

Transport and disorder in Dirac materials

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Henry Legg

aus

Jersey

Channel Islands



Köln, 2019

Berichterstatter: Prof. Dr. Achim Rosch
Priv.-Doz. Michael Scherer
Prof. Dr. Simon Trebst

Tag der mündlichen Prüfung: 29. November 2019

Tá mo cheannsa liath le bliain
's ní le críonnacht é

Abstract

The Dirac equation was able to unite relativity with quantum mechanics and successfully describe the behaviour of spin- $\frac{1}{2}$ particles. Over the past few decades it has been discovered that the electronic properties of many materials is also governed by emergent physics akin to that described by the Dirac equation, these materials are known as “Dirac materials”. Often there is a deep connection between topology and the appearance of linearly dispersing electronic bands which result in a material’s Dirac-like physics.

In this thesis we will investigate the impact disorder has on Dirac materials. In particular we will be interested in the theoretical description of transport properties – such as the electrical conductivity of a material – that result from their Dirac physics. Dirac materials provide a fascinating playground for the study of novel quantum mechanical phenomena, both theoretically and in the lab. As such, many of the examples in this thesis are the product of close theoretical and experimental collaborations.

We begin this thesis with a detailed overview of the ever-growing class of materials which obey a Dirac-like description and by introducing many of the concepts used in later chapters. Having done this we turn to a discussion of disorder. Of particular importance will be that Dirac electrons are protected from back-scattering off impurity potentials that retain the symmetries protecting the Dirac point. We will use our knowledge of disordered Dirac materials to calculate the conductivity of the surface of a topological insulator.

In the second half of this thesis we will discuss three novel phenomena which we theoretically describe and have been experimentally observed in Dirac materials: Firstly, we will discuss how it is possible to enable back-scattering in a Dirac material, in a controlled manner. We will see that this is achieved by the application of a magnetic field in the plane of a topological insulator’s surface which leads to an anisotropy of magnetoresistance and, associated to this, a planar Hall effect. Secondly, we will discuss confinement of Dirac surface states on a very thin nanowire. We will show that the quantisation of the wave-function around the wire leads to oscillatory behaviour of the resistivity that has also been experimentally observed. Finally, we turn to 3d Dirac semi-metals, we will show that their quasi-1d physics in a strong magnetic field leads to a magnetoresistivity that is strongly dependent on the angle of the applied magnetic field when there are multiple Fermi-surfaces in the Brillouin zone.

Kurzzusammenfassung

Die Dirac-Gleichung konnte die Relativitätstheorie mit der Quantenmechanik vereinen und das Verhalten von Spin- $\frac{1}{2}$ -Partikeln erfolgreich beschreiben. In den letzten Jahrzehnten wurde festgestellt, dass die elektronischen Eigenschaften vieler Materialien durch emergente Freiheitsgrade bestimmt werden, ähnlich denjenigen, die durch die Dirac-Gleichung beschrieben werden. Diese Materialien sind bekannt als „Dirac-Materialien“. Häufig besteht ein tiefer Zusammenhang zwischen der Topologie und dem Auftreten elektronischer Bänder mit linearer Dispersion, die zu einer Dirac-ähnlichen Physik eines Materials führen.

In dieser Arbeit werden wir untersuchen, welche Auswirkungen Unordnung auf Dirac Materialien hat. Insbesondere werden wir uns mit der theoretischen Beschreibung von Transporteigenschaften - wie z.B. die elektrische Leitfähigkeit eines Materials - befassen, die sich aus der Dirac-Physik ergeben. Dirac-Materialien bieten einen faszinierenden Spielwiese für die Erforschung neuer quantenmechanischer Phänomene, sowohl theoretisch als auch im Labor. Aus diesem Grund sind viele der Ergebnisse in dieser Arbeit das Produkt enger theoretischer und experimenteller Zusammenarbeit.

Wir beginnen diese Arbeit mit einem detaillierten Überblick über die ständig wachsende Klasse von Materialien, auf die eine Dirac-ähnliche Beschreibung zutrifft und stellen viele in nachfolgenden Kapiteln verwendete Konzepte vor. Anschließend wenden wir uns einer Diskussion über Unordnung zu. Von besonderer Bedeutung wird sein, dass Dirac-Elektronen vor Rückstreuung an Verunreinigungspotentialen geschützt sind, die die Symmetrien zum Schutz des Dirac-Punktes beibehalten. Wir werden unser Wissen über ungeordnete Dirac-Materialien nutzen, um die Leitfähigkeit der Oberfläche eines topologischen Isolators zu berechnen.

In der zweiten Hälfte dieser Arbeit werden wir drei neue Phänomene behandeln, die wir theoretisch beschreiben und die experimentell in Dirac-Materialien beobachtet worden sind: Zunächst werden wir diskutieren, wie es möglich ist, die Rückstreuung in einem Dirac-Material kontrolliert zu ermöglichen. Wir werden sehen, dass dies durch die Anwendung eines Magnetfeldes in der Ebene der Oberfläche eines topologischen Isolators erreicht wird, was zu einem anisotropen magnetoresistiven Effekt und damit verbunden zu einem planaren Hall-Effekt führt. Zweitens werden wir den Einschluss von Dirac-Oberflächenzuständen auf sehr dünnen Nanodrähten untersuchen. Wir werden zeigen, dass die Quantisierung der Wellenfunktion um den Draht zu Oszillationen des spezifischen Widerstands führt, das auch experimentell beobachtet wurde. Schließlich wenden wir uns 3D-Dirac-Halbmatalen zu; wir werden zeigen, dass ihre quasi-1d-Physik in einem starken Magnetfeld zu einem magnetoresistiven Effekt führt, der stark vom Winkel des angelegten Magnetfeldes abhängig ist, wenn es mehrere Fermioberflächen in der Brillouinzone gibt.

Table of contents

Introduction	1
1 Dirac Materials	5
1.1 The Dirac equation	5
1.2 Graphene	8
1.2.1 Tight binding model	8
1.2.2 Emergent Dirac physics	9
1.2.3 Generating a mass	11
1.3 The Haldane model	12
1.3.1 Breaking inversion symmetry	12
1.3.2 Breaking time-reversal symmetry	13
1.3.3 Chern number	14
1.3.4 Edge state	15
1.4 Topological insulators in 2d	17
1.4.1 The Kane-Mele Model	17
1.4.2 \mathbb{Z}_2 topological insulators	18
1.4.3 Realisations of 2d TIs	20
1.5 Topological insulators in 3d	21
1.5.1 Topological invariant(s)	21
1.5.2 Surface state Hamiltonian	22
1.5.3 Realisations of 3d TIs	24
1.6 Dirac and Weyl semi-metals in 3d	26
1.6.1 Dirac semi-metals	26
1.6.2 Weyl semi-metals	27
1.6.3 Surface states	29
1.7 Other Dirac systems	29
1.7.1 Topological crystalline insulators	30
1.7.2 d-wave superconductors	30
1.7.3 Majorana metals	30
2 Disorder in Dirac materials	33
2.1 Disordered 3d topological insulator surface	34
2.1.1 The T-matrix	35

Table of contents

2.2	Impurity averaging	37
2.2.1	Kohn-Luttinger expansion	39
2.3	The self-energy	40
2.3.1	Self-consistent T-matrix	41
2.3.2	Self-consistent Born approximation	42
2.3.3	Self-consistent Unitary approximation	43
2.4	Density of states	43
3	Quasi-particle interference from the surface of topological insulators	45
3.1	LDOS due to an impurity	45
3.1.1	LDOS oscillations on a TI surface	46
3.1.2	Point-like scatterer	47
3.1.3	Long-range scatterers	49
3.1.4	Effect of quasi-particle broadening	49
3.1.5	Magnetic impurities	50
3.1.6	QPI due to massive Dirac fermions	50
3.2	QPI from $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ surfaces	51
4	Transport in Dirac matter	55
4.1	Semi-classical transport	55
4.1.1	The Boltzmann equation	56
4.1.2	The transport scattering time	57
4.2	Linear response	59
4.2.1	Bare bubble	61
4.2.2	Vertex corrections	62
4.3	Tuning the Fermi-energy in real materials	64
4.3.1	Compensation of the bulk	65
4.3.2	Surface gating	65
5	Planar Hall effect from the surface of topological insulators	69
5.1	Lifting topological protection	69
5.2	Scattering from magnetised resonances	72
5.2.1	Surface state self-energy	73
5.2.2	Density of states	75
5.2.3	Finite magnetic field	76
5.3	Conductivity in a parallel field	76
5.3.1	Vertex corrections	77
5.4	Anisotropic MR and planar Hall effect in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$	80
5.4.1	The planar Hall effect	80
5.4.2	Surface characterisation using the PHE	81

6	Quantum confinement in topological insulator nanowires	85
6.1	Surface states on a wire	85
6.1.1	A topological insulator wrapped up	85
6.1.2	Density of states and charge density	87
6.2	A disordered TI nanowire	88
6.2.1	T-matrix approximation	88
6.2.2	Conductivity	89
6.2.3	Dilution of impurity strength	90
6.2.4	Matrix elements between sub-bands	91
6.3	Quantum confinement in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ nanowires	92
6.3.1	Bulk insulating nanowires	93
6.3.2	Resistance of real nanowires	94
7	Disordered Dirac materials in a magnetic field	99
7.1	Landau levels in 3D Dirac materials	99
7.2	Disordered Landau levels	102
7.2.1	Disordered 2d Landau levels	103
7.2.2	Disordered 3d Landau levels	105
7.2.3	Friedel oscillations in Dirac semi-metals in the quantum limit	106
7.3	Transport of semi-metals in a magnetic field	108
7.3.1	Chiral anomaly	108
7.3.2	Quantum oscillations in semi-metals	109
8	Transport spectroscopy of 3D Dirac materials in the quantum limit	111
8.1	Angular dependent magnetoresistance in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$	111
8.2	Geometry in a magnetic field	113
8.3	Angular dependent scattering	114
8.3.1	Long range impurities	114
8.3.2	Interaction corrections	116
8.3.3	Resistivity in the quantum limit	118
	Conclusions and outlook	121
	References	123
	Appendix A Quasi-particle interference from massive Dirac fermions	131
	Appendix B Matsubara sum and vertex corrections	133
	Appendix C Parameter averaging in the planar Hall effect	137
	Appendix D 3d Dirac Landau level wave-functions	139

Introduction

The concept of an electron band, resulting from the underlying periodic lattice structure of a solid, was first pioneered by Bloch and is potentially the most important concept in condensed matter physics. On its simplest level, band theory enables us to explain why some materials are metallic – such as copper – and some materials insulators – such as quartz –, all we have to do is fill up the bands with the available electrons and see whether those bands have electronic states remaining at the Fermi-level – leading to a metal – or whether they are completely full – leading to an insulator. In the second half of the last century an extensive research focus was put on materials that behave in a manner somewhere between a metal and an insulator. Such materials are known as semi-conductors and have bands which are either only *just* occupied or only *just* empty. The most prominent example is of course silicon, a material that most of our modern technology based economy is built upon. From a practical perspective we are now reaching the limit of how far we can push these semi-conductors. For instance, the latest computer processor architectures are as small as just a few dozen atoms in length and nearing the limits of physical viability.

The answer to these practical questions will almost certainly lie in new material types. Currently leading the race to replace silicon is graphene, a single layer of carbon arranged in a hexagonal lattice that was first produced in 2005 [1]. The band structure of graphene close to the Fermi-level consists – as we will see in the next chapter – of two bands that touch at a single point. Due to its band structure, graphene has an extremely high electron mobility and heat conductivity, making it an excellent candidate for the post-silicon processors of the future.

Yet the practical advantages of graphene go hand-in-hand with many fascinating theoretical questions. Despite the undeniable usefulness of semi-conductors, from a theoretical perspective they can largely be thought of as simply poor metals. Many of their electronic features can still be understood just by considering a model of a low density electron gas. This means that electrons in both semi-conductors and normal metals are well described by a quadratic relation between momentum and energy, as found in the Schrödinger equation. In stark contrast, the band “touching” in graphene leads to a linear dependence of energy and momentum equivalent to a (massless) Dirac equation better known to high-energy physics.

Graphene is definitely not the only material with low energy excitations described by a Dirac-*like* equation. Another instance can be found on the edges and surfaces of 2d and 3d topological insulators – materials that, as will be explained in the next chapter, act

like an insulator in their bulk but have semi-metallic Dirac-*like* states on their surface. A further example is the recent realisation of 3d materials with bulk bands crossings at a single point. The abundance of materials predicted and experimentally realised has lead to the definition of a new *class* of materials, the **Dirac materials** [2], all of which have a linear dispersions at low energy that touch at a point (or a set of points) in the Brillouin zone and are well described by a Dirac-*like* equation.

In the same way semi-conductors were a fillip to technological innovation in the last century, it is not unreasonable to expect that Dirac materials will have many real-world applications in the coming decades. That is not, however, the only reason to be interested in this class of materials. The access to Dirac fermions in a condensed matter setting also provides the opportunity to probe more fundamental questions about Nature, without the need for large colliders. For instance the chiral anomaly – the non-conservation of chiral charge, see next chapter – was predicted to occur in certain 3d Dirac materials [3] and has (potentially) been observed in real materials [4]. Such anomalies describe how quantum effects can violate symmetries and play an important role in cosmology and high-energy physics, providing (at least in part) answers to the question why there is more matter than antimatter in our universe. Furthermore, the existence of a Dirac-*like* dispersion is closely related to topology, in particular the physics of Berry phases that a wave function picks up when it moves in a closed loop in momentum space. One consequence of the Berry phase in Dirac matter that will be extremely important for our purposes is that it provides a “topological protection” of the Dirac electrons, meaning that electrons in Dirac materials are prohibited from backscattering in the absence of symmetry breaking perturbations.

This thesis will be focussed on the transport properties of Dirac matter and the closely related physics of disorder, which is always present in real samples. Measuring the conductivity of a material is one of the most fundamental experiments that can be performed, yet as we shall see, not only is there a unique transport phenomenology associated with Dirac materials but such measurements by themselves can reveal much of their fundamental properties; for instance in Chapter 4 we will show how the above mentioned topological protection can be lifted in a controlled and measurable manner.

A significant proportion of this thesis is directly in collaboration with experimental work, showing the direct relevance of the theoretical concepts developed here to real-world applications. However, this also poses challenges since combining theory and experiment often requires a more thorough analysis of both and, whilst providing answers, often also leads to new questions and surprises.

This thesis is written in a format which (hopefully) presents concepts in a logical progression, such that each chapter is built on the foundations laid in preceding chapters. As a result most chapters are not based on a single project; when they are this is indicated at the start of the chapter. The layout is as follows: **Chapter 1** will develop the concept of Dirac materials as a class, providing a broad overview of both theoretical predictions and experimental realisations. discussing some of their fascinating properties. **Chapter 2**

provides a largely pedagogical introduction to the influence of disorder, using the surface of a topological insulator as a working example. In particular we introduce the T-matrix, which we will make use of in our discussions of phenomena in later chapters. **Chapters 3, 4, 5, and 6** will discuss specific theoretical effects in 3d topological insulators (TIs), all of which have also been experimentally observed. In particular, **chapter 3** links in with the discussion of disorder in chapter 2, showing that a Dirac electron's protection from backscattering can be directly measured by quasi-particle interference using scanning tunnelling spectroscopy. **Chapter 4** provides a basis for the next 2 chapters by introducing transport and calculating the conductivity of a disorder topological insulator surface. In **chapter 5** we provide a method – using an in-plane magnetic field – that lifts this protection whilst retaining the Dirac physics of the topological insulator surface. In **chapter 6** our discussion moves to thin wires of TI; we will show that the quantisation of the TIs surface state around such a wire leads to experimentally measurable oscillations in resistivity as a function of gate voltage. In **chapter 7 and 8** we then turn to magnetoresistance properties of three dimensional Dirac materials. **Chapter 7** is again largely pedagogical in nature, presenting the unique response of a Dirac material to a magnetic field and some of the resulting phenomena. In **chapter 8** we will discuss a specific effect, namely the enhancement of magnetoresistance in materials with multiple Dirac points in their Brillouin zone. At the end we provide an overview and discuss avenues for future research. Several auxiliary calculations are relegated to the appendices.

Dirac Materials

In the introduction we outlined the definition of Dirac materials as a *class* of materials with low energy excitations well described by a Hamiltonian akin to the Dirac equation. In this chapter we clarify that definition first by discussing the Dirac equation in its general context as the description of spin- $\frac{1}{2}$ relativistic fermions. We then turn to systems where the Dirac equation appears in a condensed matter setting. Our first example will be **graphene** [1, 5, 6], a single layer of carbon arranged in a honeycomb lattice. Having established its Dirac physics we then proceed to break it and discuss the **Haldane model** [7], although in a sense only half a Dirac material, it provides us with a concrete example to develop many of the concepts of topology that we will require to understand other members of the Dirac material class. Most notably among these will be the **topological insulators**, first predicted by Kane and Mele in graphene [8, 9] and then extended to other 2d materials and later also to 3d [10–16]. Since these topological insulators and experiments on them are central to this thesis we also discuss their experimental realisations [17–19]. We then move to another major subset of the Dirac materials, the **Dirac** and **Weyl semi-metals** [20, 21, 2, 22–25]. These also feature prominently in later chapters and we will once again discuss both their theoretical foundations and experimental realisations, as well as their novel surface states. In the penultimate subsection we collect several members of the Dirac materials class that, while not of immediate relevance to this thesis, nonetheless show the breadth of Dirac realisations [26–32].

Our central purpose in this chapter is simply to develop an understanding of Dirac materials that we can draw on in later chapters, in particular in our discussion of several phenomena that they exhibit. Nonetheless, since there is also a beautiful mathematical structure underlying the topological features of Dirac materials, this chapter will provide the salient features of that framework. Many excellent reviews on graphene [6], topological insulators [14–16], and Dirac materials as a whole [2, 22, 23], already exist and can be referred to for more detailed discussions. While this chapter is designed to give a taste of the huge variety of Dirac materials that exist in Nature it is far from a comprehensive list; indeed the number of members of the Dirac material class has seen a substantial growth in the last two decades, a growth that shows no sign of abating.

1.1 The Dirac equation

Although the Schrödinger equation’s simplicity makes it a good teaching tool, the Dirac equation for spin- $\frac{1}{2}$ relativistic particles is – arguably – the real cornerstone of quantum mechanics [33]. Dirac’s accomplishment was not only being able to successfully combine

Chapter 1. Dirac Materials

relativity and quantum mechanics, his equation also provided a theoretical explanation for spin and using it he predicted the existence of anti-matter [34], only discovered experimentally *after* Dirac published his seminal paper [35]. The Dirac equation is a 4×4 matrix equation in 3 spatial and 1 space-time dimension. It can be written in several equivalent forms, one of the most elegant being

$$(i\hbar\gamma^\mu\partial_\mu - mc\mathbb{1}_4)\psi = 0, \quad (1.1)$$

where a sum over the index $\mu = 0, 1, 2, 3$ is implied and we made use of the covariant 4-gradient $\partial_\mu = (\frac{1}{c}\partial_t, \partial_x, \partial_y, \partial_z)$. Most importantly the objects γ^μ are 4×4 matrices satisfying the anti-commutation relations $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{1}_4$ – the Dirac algebra¹ – (where $\eta^{\mu\nu}$ is the Minkowski metric²) and the wave-function ψ is a 4-component bi-spinor.³ The Dirac equation is first order in both space and time putting them on an equal footing, unlike the Schrödinger equation which is only first order in time. In particular, since the matrices γ^μ transform contravariantly and ∂_μ covariantly the equation as a whole is invariant under Lorentz transformations and so consistent with special relativity. A particularly simple basis for the γ -matrices can be formulated in terms of the 2×2 Pauli-matrices σ^i such that

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix}, \quad \& \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \quad (1.2)$$

Let us now find the solutions of the Dirac equation for a free particle, which without loss of generality we take to move along the z -direction with momentum p_z and have energy ε . By translational invariance it should be clear that solutions will be plane waves of the form $\psi = (\mathbf{u}_1, \mathbf{u}_2)^T e^{-i(\varepsilon t - p_z z/c)/\hbar}$, where $\mathbf{u}_{1,2}$ are the two 2-spinor components of the full bi-spinor. Inserting this ansatz wave-function to the Dirac equation, we find the spinors $\mathbf{u}_{1,2}$ must satisfy

$$\begin{pmatrix} (\varepsilon - mc^2)\mathbb{1} & -\hbar\sigma_z p_z \\ \hbar\sigma_z p_z & -(\varepsilon + mc^2)\mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} = 0 \quad \text{or equivalently} \quad \mathbf{u}_{1,2} = \frac{\hbar p_z \sigma_z}{\varepsilon \pm mc^2} \mathbf{u}_{2,1}. \quad (1.3)$$

The corresponding eigenvalues are $\varepsilon = \pm\sqrt{\hbar^2 p_z^2 + m^2 c^4}$ – the energy formula for relativistic particles – and we can choose a basis for the four (unnormalised) eigenstates such that

$$\psi_\uparrow^+ = \begin{pmatrix} 1 \\ 0 \\ \frac{\hbar p_z}{\varepsilon + mc^2} \\ 0 \end{pmatrix}, \quad \psi_\downarrow^+ = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{-\hbar p_z}{\varepsilon + mc^2} \end{pmatrix}, \quad \psi_\uparrow^- = \begin{pmatrix} \frac{-\hbar p_z}{-\varepsilon + mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \& \quad \psi_\downarrow^- = \begin{pmatrix} 0 \\ \frac{\hbar p_z}{-\varepsilon + mc^2} \\ 0 \\ 1 \end{pmatrix}. \quad (1.4)$$

¹The Dirac algebra is simply a special case of the Clifford algebra in 3+1 dimensions.

²In what follows we will use $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$ but $\eta^{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ is also possible. Since some people have strong feelings about which definition is best, I apologise if this is not your preference.

³A spinor is a mathematical object that changes sign when rotated by 360° .

1.1. The Dirac equation

These eigenstates consist of two positive and two negative energy solutions – indicated by \pm superscript –, positive energies correspond to matter and negative to anti-matter. Additionally each element within a pair of matter or antimatter states has opposite spin – indicated by \uparrow or \downarrow .

In a condensed matter setting we will be interested in situations where $\hbar|p_z| \gg mc^2$, known as the *ultra-relativistic* limit [22]. At first sight this might look like a very high energy situation, in fact this occurs in a Dirac material because symmetries force the (effective) mass of the electrons to be zero (or very small, if weakly broken). Within this limit the eigenstates are said to become *chiral*. That is to say that they become simultaneous eigenstates of the helicity operator $\hat{h} = \frac{\mathbf{S} \cdot \mathbf{p}}{|\mathbf{p}|}$, which is not Lorentz invariant, and the left- and right-handed chirality operators $\hat{C}_{L/R} = \frac{\hbar}{2}(\mathbb{1} \mp \gamma_5)$, which are Lorentz invariant. Here the components of the spin \mathbf{S} are defined as

$$\hat{S}_i = \frac{\hbar}{2} \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \quad \text{and the matrix} \quad \gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix}. \quad (1.5)$$

The helicity measures the spin-component in the direction of the particle’s momentum – in our example above, the z -direction. We actually chose a basis above so that, regardless of the mass, the eigenstates in Eq. (1.4) are always eigenstates of the helicity operator. In particular, for $p_z > 0$, the eigenstates ψ_{\uparrow}^{\pm} have helicity $+\frac{\hbar}{2}$ and ψ_{\downarrow}^{\pm} have helicity $-\frac{\hbar}{2}$. When the mass vanishes the energy of a state is simply $\varepsilon = \pm|p_z|$ and we see – as claimed – in this limit our eigenstates from Eq. 1.4 are simultaneous eigenstates of the chirality operator with a left or right handed chirality equal in magnitude to their helicity. When the mass is finite, however, the eigenstates of helicity and chirality are no longer the same.

To understand why helicity and chirality eigenstates are not the same for a finite mass, we start by considering a basis in terms of eigenstates of chirality, these are simply left-handed state $\psi_L = (\mathbf{u}_L, -\mathbf{u}_L)^T e^{-i(\varepsilon t - p_z z/c)/\hbar}$ and right hand state $\psi_R = (\mathbf{u}_R, \mathbf{u}_R)^T e^{-i(\varepsilon t - p_z z/c)/\hbar}$, where in terms of our spinor basis above $\mathbf{u}_L = \frac{1}{2}(\mathbf{u}_1 - \mathbf{u}_2)$ and $\mathbf{u}_R = \frac{1}{2}(\mathbf{u}_1 + \mathbf{u}_2)$. In particular in order to solve the Dirac equation the chiral spinors $\mathbf{u}_{L/R}$ must satisfy

$$(\varepsilon + \hbar \boldsymbol{\sigma} \cdot \mathbf{p}) \mathbf{u}_L = mc^2 \mathbf{u}_R \quad \text{and} \quad (\varepsilon - \hbar \boldsymbol{\sigma} \cdot \mathbf{p}) \mathbf{u}_R = mc^2 \mathbf{u}_L. \quad (1.6)$$

It is now obvious that, for a finite mass, the left-handed and right-handed chiral states are coupled in the Dirac equation by the mass-energy mc^2 . The reason that the helicity and chirality eigenstates match in the ultra-relativistic limit, $m = 0$, is because this coupling vanishes and the Dirac equation becomes diagonal in terms of the chiral spinors $\mathbf{u}_{L/R}$. This special case of the massless Dirac equation is called the *Weyl* equation, which we will return to later in the chapter. Of course we could have also understood all of this from a relativistic perspective since the only way a Lorentz invariant quantity like chirality could correspond to a non-Lorentz invariant quantity like helicity is if the particle’s mass is zero such that there is no frame of reference where we can “overtake” the particle.

We have discovered several features of the Dirac equation that we will see transfer to a condensed matter setting: 1) At high energies, or low mass, the eigenenergies are linear in the momentum $|\mathbf{p}|$. 2) The equation is a matrix equation where the basis matrices satisfy the Dirac algebra. 3) For each positive energy eigenstate a corresponding negative energy anti-matter eigenstate exists, in a condensed matter context this is known as *particle-hole* symmetry.⁴ 4) In the case of massless fermions the spin and momentum are *locked* such that chirality and helicity are the same quantity, we will refer to this as *spin-momentum locking*. 5) For massive fermions the eigenstates of chirality no longer solve the Dirac equation, we will see that this is related to whether or not spin-flip scattering is allowed.

1.2 Graphene

Having established the key features of the Dirac equation we now proceed to show how this can be achieved in a condensed matter setting. The archetypal Dirac material is *graphene* [1, 5, 6]. Graphene consists of a single 2d layer of carbon atoms arranged in a hexagonal lattice (see Fig. 1.1). It is actually this hexagonal lattice that is key and many other 2d hexagonal materials like silicene, germanene, and others [36–38] have very similar dispersions.⁵ The hexagonal lattice is bipartite, consisting of two triangular sublattices – we will label these A and B , they are indicated by different colours in Fig. 1.1.

1.2.1 Tight binding model

To a good approximation the only free electrons in graphene belong to the half-filled p_z orbital, with z perpendicular to the lattice plane [6]. Since the largest overlap of these orbitals is with their nearest neighbours and this overlap is isotropic because z lies normal to the plane, it is natural to assume an isotropic nearest neighbour tight-binding model

$$H_G = t \sum_{\mathbf{r}_A} \sum_{i=1,2,3} \left(c_B^\dagger(\mathbf{r}_A + \boldsymbol{\delta}_i) c_A(\mathbf{r}_A) + c_A^\dagger(\mathbf{r}_A) c_B(\mathbf{r}_A + \boldsymbol{\delta}_i) \right), \quad (1.7)$$

where c_A (c_A^\dagger) and c_B (c_B^\dagger) are annihilation (creation) operators on the A and B sublattices respectively; $\mathbf{r}_A = a \frac{\sqrt{3}}{2} (m \hat{\mathbf{x}} + (2n + m) \hat{\mathbf{y}})$, with $n, m \in \mathbb{Z}$ are the lattice sites of the triangular A sublattice; and the nearest neighbour vectors $\boldsymbol{\delta}_i$ are defined as $\boldsymbol{\delta}_1 = a(1, 0)$, $\boldsymbol{\delta}_2 = \frac{a}{2}(-1, \sqrt{3})$, and $\boldsymbol{\delta}_3 = \frac{a}{2}(-1, -\sqrt{3})$, where a is the lattice constant. These lattice vectors are shown in Fig. 1.1. The Hamiltonian Eq. (1.7) describes hopping of electrons from A to adjacent B sites and vice versa.

⁴Most condensed matter Dirac systems are actually only approximately particle-hole symmetric. For example we will see the bulk bands can cause a particle-hole asymmetry in 3d topological insulators.

⁵The key difference between all these materials is the size of the spin-orbit coupling, which is a perturbation that opens a band-gap at the Dirac points. We will see later that this perturbation is actually incredibly important in determining the hexagonal lattice's electronic properties.

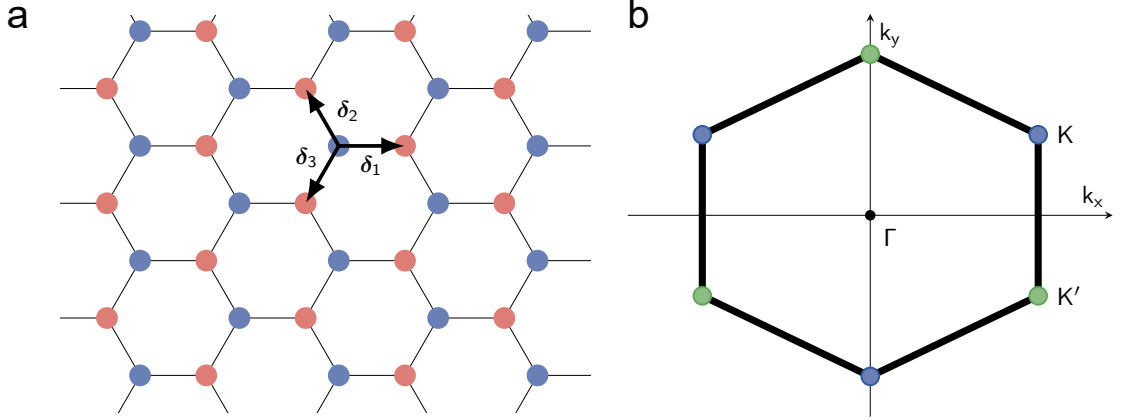


Fig. 1.1 **Hexagonal lattice in real space and its Brillouin zone:** a) The hexagonal lattice of carbon atoms that form graphene. The lattice is bipartite with each sublattice indicated by blue or red. The lattice vectors δ_1 , δ_2 , and δ_3 are indicated. b) The first Brillouin zone of the hexagonal lattice. For the nearest neighbour tight binding model Dirac nodes live at the zone corners, the K and K' points (labelled and coloured green, blue respectively).

To diagonalise the Hamiltonian Eq. (1.7) we Fourier transform the operators $c_A(\mathbf{r}) \rightarrow a_{\mathbf{k}}$ and $c_B(\mathbf{r}) \rightarrow b_{\mathbf{k}}$ giving

$$H = t \sum_{\mathbf{k} \in \text{BZ}} \begin{pmatrix} a_{\mathbf{k}}^\dagger & b_{\mathbf{k}}^\dagger \end{pmatrix} \begin{pmatrix} 0 & f_{\mathbf{k}}^* \\ f_{\mathbf{k}} & 0 \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}} \\ b_{\mathbf{k}} \end{pmatrix}, \quad \text{where} \quad f_{\mathbf{k}} = \sum_{n=1,2,3} e^{i\delta_n \cdot \mathbf{k}}. \quad (1.8)$$

The eigenenergies of the Hamiltonian can now be found by diagonalising it, they are simply $\varepsilon_{\mathbf{k}} = \pm |f_{\mathbf{k}}|$ defining a conduction band (+) which is always positive in energy and valence band (−) which is always negative in energy. The most important feature of the dispersion $\varepsilon_{\mathbf{k}}$, shown in Fig. 1.2, are the positions $\mathbf{K} = \frac{2\pi}{3} \left(2, \frac{1}{\sqrt{3}} \right)$ and $\mathbf{K}' = \frac{2\pi}{3} \left(2, -\frac{1}{\sqrt{3}} \right)$ where the bands *just* touch at a point. These K-points correspond to the Brillouin zone corners as shown in Fig. 1.1 and the energy vanishes here because the three contributions to $f_{\mathbf{k}}$ perfectly destructively interfere.

1.2.2 Emergent Dirac physics

Since the bands in graphene are naturally half-filled, excitation energies are defined with respect to the band crossing which we define to be $\varepsilon = 0$. In particular to obtain the low-energy excitations we simply need to Taylor expand about the K-points using $\mathbf{k} = \tau \mathbf{K} + \mathbf{q}$, where we use $\tau = \pm 1$ since this maps \mathbf{K} to \mathbf{K}' when negative. After expanding the Hamiltonian we see that it has a distinctly Dirac-like form

$$H_{\tau \mathbf{K}} = -\frac{3at}{2} \sum_{\mathbf{k} \in \text{BZ}, \tau} \begin{pmatrix} a_{\mathbf{q}, \tau \mathbf{K}}^\dagger & b_{\mathbf{q}, \tau \mathbf{K}}^\dagger \end{pmatrix} \begin{pmatrix} 0 & q_x - i\tau q_y \\ q_x + i\tau q_y & 0 \end{pmatrix} \begin{pmatrix} a_{\mathbf{q}, \tau \mathbf{K}} \\ b_{\mathbf{q}, \tau \mathbf{K}} \end{pmatrix}, \quad (1.9)$$

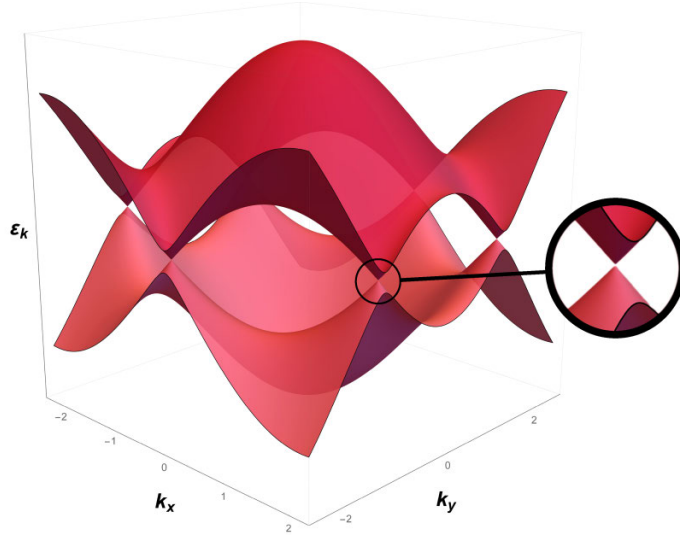


Fig. 1.2 **Energy dispersion in graphene:** The relation, $\varepsilon_{\mathbf{k}}$, between energy and momentum for nearest neighbour hopping in graphene. At the Brillouin zone corners, the K-points, the conduction and valence bands touch. The low energy excitations about the K-points are well described by a massless 2d Dirac equation.

where we have used the fact that we are entirely free to multiply our creation operators by any arbitrary phase leaving a phase difference between A and B sublattices. The full 4×4 Hamiltonian can then be brought into the form of two copies of what we will see is the 2d Dirac equation⁶

$$H_G = v_F \sum_{\mathbf{q}} (\mathbf{a}_{\mathbf{q},\mathbf{K}}^\dagger, \mathbf{a}_{\mathbf{q},-\mathbf{K}}^\dagger) \begin{pmatrix} (\boldsymbol{\sigma} \cdot \mathbf{q}) & 0 \\ 0 & (\boldsymbol{\sigma} \cdot \mathbf{q})^* \end{pmatrix} \begin{pmatrix} \mathbf{a}_{\mathbf{q},\mathbf{K}} \\ \mathbf{a}_{\mathbf{q},-\mathbf{K}} \end{pmatrix}, \quad (1.10)$$

where $\mathbf{a}_{\mathbf{q},\pm\mathbf{K}} = (a_{\mathbf{q},\pm\mathbf{K}}, b_{\mathbf{q},\pm\mathbf{K}})$ is a 2-spinor of the creation operators on each sub-lattice and for each K-point. The σ_i are simply the Pauli-matrices and referred to as the *pseudo-spin* given their analogy with actual spins. The τ index is called the *valley* degree of freedom. It is important to note that there is also double degeneracy due to real spin in all of these equations, in particular the pseudospin that appears in our Dirac Hamiltonian is *not* the real spin; this is not the case for all Dirac materials as we shall see when we discuss topological insulators.

There are some key differences between our Hamiltonian Eq. 1.10 and the Dirac equation of Eq. 1.1: First, the Dirac equation is in 3+1 dimensions whereas our Hamiltonian is explicitly an equation in 2 spatial dimensions. Second, we are limited by a length scale, the lattice constant a , such that our linear Hamiltonian is only valid for low energies or, equivalently, small momenta $|\mathbf{q}| \ll \frac{1}{a}$. Nonetheless, if we now compare these two copies of the low-energy Hamiltonian $H_{\pm\mathbf{K}} = v_F \boldsymbol{\sigma} \cdot \mathbf{q}$, to the original Dirac equation we see it is the 2d analog and the important features from the end of the last section are all present: 1)

⁶This description ignores that the bands are actually connected in the high energy regime and so is only valid for small excitations around the Dirac point.

The eigenenergies $\varepsilon_{\mathbf{q}} = \pm v_F |\mathbf{q}|$ are linear in momentum \mathbf{q} with the speed of light replaced by the Fermi-velocity $v_F = -\frac{3at}{2}$. 2) There is an underlying matrix structure that obeys the anticommutation relation $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$, the 2d Dirac algebra. 3) Every positive energy state has an equivalent negative energy state. Unlike the Dirac equation, where negative energy states indicated the existence of anti-matter, the physical interpretation is more clear in graphene as they simply correspond to removing a hole from the occupied states in the valence band. 4) There is a locking of *pseudospin* and momentum within each valley of the form $\mathbf{q} \cdot \boldsymbol{\sigma}$. 5) We can add a mass by introducing a term of the form $m\sigma_z$ in pseudospin space that acts as mass, the physics of which we now discuss in more detail.

1.2.3 Generating a mass

To end our discussion on graphene we need to consider what symmetries were at play that forced us to have gapless Dirac points at the corners of the Brillouin zone. In particular two symmetries were key [22]: Inversion symmetry \mathcal{P} and time reversal symmetry \mathcal{T} .

Inversion maps one sublattice site to the opposite and one valley to the other since under its action $\mathbf{K} \rightarrow -\mathbf{K}$. Within our Pauli basis we therefore see that the inversion operator is given by $\mathcal{P} = \tau_1 \otimes \sigma_x$, with \otimes the Kronecker product. Acting this on Eq. 1.10 we find $\mathcal{P}H(\mathbf{q})\mathcal{P}^{-1} = H(-\mathbf{q})$ such that the Hamiltonian is invariant under inversion symmetry, as expected since graphene's lattice is inversion symmetric.⁷

For spinless fermions, time reversal symmetry also maps $\mathbf{K} \rightarrow -\mathbf{K}$ and involves applying K_c , complex conjugation, unlike inversion, however, it has no impact on the pseudospin and can therefore be written $\Theta = \tau_1 \otimes \mathbb{1}_2 K_c$. Once again we find that the Hamiltonian is invariant under its action $\Theta H(\mathbf{q})\Theta^{-1} = H(-\mathbf{q})$, which is expected since we haven't applied any magnetic fields.

Now let us consider possible mass-like perturbations that will open a gap at the Dirac points. Additional terms in the Hamiltonian proportional to σ_x or σ_y will simply shift the position of the Dirac cones. In contrast, a term proportional to $\mathbb{1} \otimes \sigma_z$ would open a gap since it is diagonal in the valley index but proportional to σ_z , the only pseudospin Pauli-matrix not appearing in our Hamiltonian, Eq. 1.10. Physically such a term creates an imbalance in the potential between the A and B sublattices. Applying inversion gives $\mathcal{P}(\mathbb{1}_2 \otimes \sigma_z)\mathcal{P}^{-1} = -\mathbb{1}_2 \otimes \sigma_z$ so that the perturbation is *not* invariant under inversion symmetry. Conversely in order to allow such a term and create a mass we *must* break inversion symmetry. This perturbation is, however, invariant under time reversal meaning that we only need to break inversion for such a term to occur.

On the other hand we can consider a perturbation creating both an imbalance between sublattice and valley degrees of freedom, $\tau_3 \otimes \sigma_z$. This is *not* invariant under time-reversal $\theta(\tau_3 \otimes \sigma_z)\theta^{-1} = -\tau_3 \otimes \sigma_z$ and so can only occur if time-reversal is broken. This perturbation is, however, invariant under inversion symmetry. The only remaining mass-like

⁷This might seem confusing since our labelling of A and B sites in Fig. 1.1 is *not* inversion symmetric, however that was simply a labelling and applying inversion of course just maps carbon atoms to carbon atoms and so inversion is a true symmetry of the lattice.

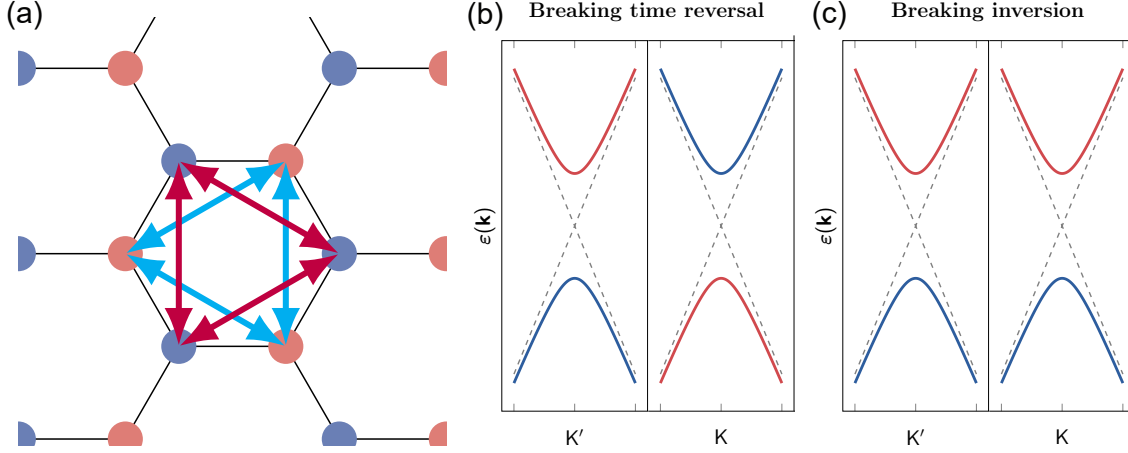


Fig. 1.3 **The Haldane model:** a) Adding next nearest neighbour hoppings to graphene, ν_1 , ν_2 , and ν_3 (see arrows), with a complex phase dependent on sublattice index breaks time-reversal symmetry and opens a gap at the Dirac points. b) The mass at the K and K' points has opposite sign for the Haldane model. c) In contrast, a sublattice dependent potential which breaks inversion symmetry causes a mass with the same sign at the K-points.

perturbations are $\tau_1 \otimes \sigma_z$ and $\tau_2 \otimes \sigma_z$ the former breaks inversion symmetry and the latter time-reversal. Therefore we see that, for spinless fermions, in order to open a gap we *must* break either inversion symmetry or time reversal symmetry and our Dirac points are protected against any perturbations that respect these two symmetries. Additionally, whether a perturbation breaks inversion or time-reversal symmetry will have a substantial impact on the *type* of insulator that occurs when a gap is opened, as we now show.

1.3 The Haldane model

The Haldane model enables us to break time-reversal symmetry in the graphene Hamiltonian Eq. (1.7) *without* a net magnetic field [7]. We shall see that the resulting insulating state is similar to that of the quantum Hall insulator and contains a chiral edge mode. The major difference between the Haldane model and a quantum Hall insulator is that the total flux through a unit cell vanishes, meaning that translational symmetry and, associated to this, Bloch's theorem still hold.⁸

1.3.1 Breaking inversion symmetry

Before we discuss Haldane's Hamiltonian let us start with the perturbation described above that retains time-reversal symmetry but breaks inversion symmetry. The most obvious

⁸The Haldane insulator itself is not really a Dirac material. Whilst its edge states are chiral they flow only in one direction – unlike the $\pm k$ of a genuine 1D Dirac system, such as the edge states of the 2D topological insulators we will see in the next section. Furthermore, although it has been realised in cold atom experiments [39], it is extremely unlikely that any real material will give rise to a single copy of the Haldane Hamiltonian. Nonetheless, it is an elegant example of the how the concept of topology can enter into a condensed matter setting and provides us with a bridge between graphene and topological insulators.

1.3. The Haldane model

way to break inversion symmetry is a staggered on-site potential [40]

$$\Delta H_{AB} = \varepsilon_A \sum_{\mathbf{r}_A} c^\dagger(\mathbf{r}_A) c(\mathbf{r}_A) + \varepsilon_B \sum_{\mathbf{r}_B} c^\dagger(\mathbf{r}_B) c(\mathbf{r}_B), \quad (1.11)$$

where $\varepsilon_A \neq \varepsilon_B$. From the last section this perturbation is of the form $M\mathbb{1}_2 \otimes \sigma_z$ in terms of the valley and pseudospin, with its “mass” given by the difference in on-site potential $M = \frac{\varepsilon_A - \varepsilon_B}{2}$. After such a perturbation has been added to graphene’s Hamiltonian, Eq. (1.7), the 2d Dirac equation in each valley takes the form $H = v_F \boldsymbol{\sigma} \cdot \mathbf{p} + M\sigma_z$ and has eigenvalues $\varepsilon_{\mathbf{k}} = \pm v_F \sqrt{\mathbf{k}^2 + M^2}$. We see that the perturbation, as expected, has opened up a gap at both K-points and that the mass M has the *same* sign at the two K-points.

1.3.2 Breaking time-reversal symmetry

We also showed at the end of the last section a perturbation $M\tau_3 \otimes \sigma_z$ breaks time-reversal but maintains inversion symmetry. It is clear from the form of such a perturbation that the mass will evidently *not* be the same at both Dirac points if we are able to break time-reversal symmetry in a manner that maintains Bloch’s theorem. Haldane’s proposal to achieve this is a staggered flux throughout the unit cell due to complex second nearest neighbour hoppings

$$\Delta H_{\text{Hal}} = t' \sum_{i=1,2,3} \left(\sum_{\mathbf{r}_A} e^{i\phi} c_A^\dagger(\mathbf{r}_A + \boldsymbol{\nu}_i) c_A(\mathbf{r}_A) + \sum_{\mathbf{r}_B} e^{-i\phi} c_B^\dagger(\mathbf{r}_B + \boldsymbol{\nu}_i) c_B(\mathbf{r}_B) \right) + \text{H.c.}, \quad (1.12)$$

where the next nearest neighbour vectors $\boldsymbol{\nu}_i$ within a unit cell are shown in Fig. 1.3 and can be obtained by combining two different nearest neighbour vectors such that the hoppings form a closed loop around a single hexagon, for example $\boldsymbol{\nu}_1 = \boldsymbol{\delta}_1 - \boldsymbol{\delta}_2$, $\boldsymbol{\nu}_2 = \boldsymbol{\delta}_2 - \boldsymbol{\delta}_3$, and $\boldsymbol{\nu}_3 = \boldsymbol{\delta}_3 - \boldsymbol{\delta}_1$ (see Fig. 1.3). The sublattice dependent phase factor $e^{\pm i\phi}$ means that, upon hopping to a next nearest neighbour, the electrons pick up a phase $\pm\phi$ depending on the direction of the hopping.

After Fourier transforming in terms of creation operators $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ on each sublattice, the full Hamiltonian is the sum of the Hamiltonian of graphene Eq. 1.13 – which only contributes to the off-diagonals – and the Fourier transform of the Haldane perturbation Eq. 1.12 – which only contributes to the diagonals – giving

$$H = t \sum_{\mathbf{k} \in \text{BZ}} \begin{pmatrix} a_{\mathbf{k}}^\dagger & b_{\mathbf{k}}^\dagger \end{pmatrix} \begin{pmatrix} \frac{t'}{t} (e^{i\phi} g_{\mathbf{k}} + e^{-i\phi} g_{\mathbf{k}}^*) & f_{\mathbf{k}}^* \\ f_{\mathbf{k}} & \frac{t'}{t} (e^{-i\phi} g_{\mathbf{k}} + e^{i\phi} g_{\mathbf{k}}^*) \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}} \\ b_{\mathbf{k}} \end{pmatrix}, \quad (1.13)$$

where $g_{\mathbf{k}} = \sum_n e^{i\boldsymbol{\nu}_n \cdot \mathbf{k}}$. In particular we can write the Haldane perturbation as

$$\Delta H_{\text{Hal}} = 2t' \sum_{i=1,2,3} (\cos(\phi) \cos(\mathbf{k} \cdot \boldsymbol{\nu}_i) \mathbb{1} + \sin(\phi) \sin(\mathbf{k} \cdot \boldsymbol{\nu}_i) \sigma_z), \quad (1.14)$$

the first term, proportional to $\mathbb{1}$, is simply a shift of the Dirac point away from zero energy and the second term, proportional to σ_z leads to a mass in the Dirac Hamiltonian. Most importantly the value of this mass is different for each Dirac point. In particular at the K- and K'-point the mass is $t' \sin(\phi) \sum_i \sin(\pm \mathbf{K} \cdot \boldsymbol{\nu}_i) = \pm t' \sin(\phi) 3\sqrt{3}$, where (+) corresponds to the K-point and (−) for the K'-point.⁹ In both cases the eigenenergies expanded about the Dirac points are $\varepsilon_{\mathbf{k}} = \pm \sqrt{v_F^2 k^2 + 27t'^2 \sin^2(\phi)}$.

1.3.3 Chern number

To understand the importance of a Chern number we must start with the concept of a *Berry phase* [41, 42].¹⁰ Consider a wavefunction $|\psi_{\mathbf{k}}\rangle$ with a (non-degenerate) energy $\varepsilon_{\mathbf{k}}$; we gradually vary \mathbf{k} such that $|\psi_{\mathbf{k}}\rangle$ always undergoes an adiabatic evolution. Now, returning to the initial position \mathbf{k} in some closed loop, Γ , surprisingly the wavefunction *does not* have to return to its initial state [44]. Rather, as Berry showed, it can pick up a global phase

$$\gamma(\Gamma) = \oint_{\Gamma} \mathbf{A}(\mathbf{k}) \cdot d\mathbf{k}, \quad (1.15)$$

where we have defined the *Berry connection* $\mathbf{A}(\mathbf{k}) = i \langle \psi_{\mathbf{k}} | \nabla_{\mathbf{k}} | \psi_{\mathbf{k}} \rangle$ which measures the phase difference between two states that are infinitesimally separated in \mathbf{k} -space. Berry showed that the total phase accumulated along the path Γ is not dependent on the specific path taken but on the topological features of the Berry connection enclosed in that path, as we will now see.

Actually we could have multiplied our state $|\psi_{\mathbf{k}}\rangle$ by some arbitrary phase $e^{i\varphi_{\mathbf{k}}}$ without any physical consequence. If we do this we see that the Berry connection is a gauge dependent object, transforming under such a transformation as $\mathbf{A}'(\mathbf{k}) = \mathbf{A}(\mathbf{k}) + \nabla_{\mathbf{k}} \varphi_{\mathbf{k}}$, just like the vector potential of a magnetic field. In fact this analogy holds and we can define the equivalent of a magnetic field, the *Berry curvature*, $\boldsymbol{\Omega}(\mathbf{k}) = \nabla_{\mathbf{k}} \times \mathbf{A}(\mathbf{k})$, which is gauge-independent and therefore a more physical object. In particular, we are interested in the integral of this Berry curvature over the Brillouin zone, which defines a non-contractible loop of \mathbf{k} since the Brillouin zone boundaries are periodic. This means the resulting integral must give us an integer multiple of 2π . The number of multiples n of 2π that are added by the closed loop around the Brillouin zone is called the *Chern* number and defined in terms of the Berry curvature as

$$n = \frac{1}{2\pi} \int_{\text{BZ}} d\mathbf{S} \cdot \boldsymbol{\Omega}_{\mathbf{k}}, \quad (1.16)$$

where \mathbf{S} surface normal of the Brillouin zone. Just as in Stoke's theorem, the important points in the Brillouin zone that contribute to the Chern number are the locations of any

⁹It should be noted that the *real* next nearest neighbour coupling $\phi = 0, \pi$ result in no mass and only a shift the Dirac points' energy. This is because such a coupling breaks neither time-reversal nor inversion symmetry and, as we showed at the end of the last section, one of these symmetries must be absent if we are to gap out the Dirac cones.

¹⁰Despite the name the “Berry” phase it was actually first discovered by Pancharatnam [43] in 1954 and then independently and with more prominence by Berry in 1984.

1.3. The Haldane model

singularities in $\Omega_{\mathbf{k}}$. We see that the Chern number is a measure of the topology of our system; it is our first example of a *topological invariant*.

Having established the existence of the Chern number, let us calculate it for both the inversion and time reversal symmetry breaking perturbations above. We saw the form of graphene's Hamiltonian resulting from one of these perturbations¹¹ in terms of $\mathbf{h} = (H_x, H_y, M)$ close to the Dirac point could be written as

$$H = H_x(\mathbf{k})\sigma_x + H_y(\mathbf{k})\sigma_y + M(\mathbf{k})\sigma_z = \mathbf{h} \cdot \boldsymbol{\sigma}. \quad (1.17)$$

Evidently any poles in the Berry curvature $\Omega_{\mathbf{k}}$ must occur at the K-points since away from the zone corners we can choose the eigenspinor of the Bloch wavefunctions as, for example, $\psi_{\mathbf{k}} = (M - |\mathbf{h}|, H_x + iH_y)^T$ without ambiguity. To calculate the contributions of the Dirac points to the Chern number we expand the vector \mathbf{h} close to them obtaining $\mathbf{h} = (v_F k_x, \tau v_F k_y, M_{\tau\mathbf{K}})$, where $\tau = \pm 1$ is our valley degree of freedom and $M_{\tau\mathbf{K}}$ is the mass at the K- or K'-point. For the inversion symmetry breaking perturbation H_{AB} the mass has the same sign at both Dirac points and so, due to the τ factor in front of k_y in \mathbf{h} , integrating any choice of eigenspinor over a closed loop around K or K' will give equal and opposite contributions to the Chern number so that $n_{AB} = 0$. In stark contrast $M_{\tau\mathbf{K}}$ had the opposite sign at each K- and K'-point for the Haldane perturbation H_{Hal} . This means when performing the equivalent closed loop integral around the K and K'-points the valley index τ and sign of the mass multiply such that the contributions to the Chern number at both K-points are the *same*. In particular the Chern number for the Haldane model is $n_{\text{Hal}} = \pm 1$, with the sign depending on the sign of the masses at the K-points. This is determined by whether the phase satisfies $0 < \phi < \pi$ or $\pi < \phi < 2\pi$.

1.3.4 Edge state

We have shown that there is a fundamental *topological* difference between the insulating state resulting from an inversion breaking perturbation such as H_{AB} and Haldane's insulating state with non-real hopping which has Chern number $n = \pm 1$. This topological difference has a physical consequence, namely, a chiral edge mode at the interface of the Haldane insulator and any normal, non-topological insulator – for instance the vacuum.

We can model such an interface by a spatially varying mass $M(y)$, where we choose the mass to change from the Haldane case $M(y > 0) = M_H \tau_z \otimes \sigma_z$ to the trivial case $M(y < 0) = M_{\text{vac}} \sigma_z \rightarrow +\infty \sigma_z$ at the interface along the line $y = 0$. For low energies $\varepsilon \ll v_F/a$, the corresponding eigenenergy equation for the spinor $\psi_{k_x}(y) = e^{-ik_x x} \varphi(y)$ from the Hamiltonian in a single valley, Eq. (1.13), has the form [22],

$$\left(v_F \sigma_x k_x - i v_F \tau \sigma_y \partial_y + \tau M(y) \sigma_z \right) \psi_{k_x}(y) = \varepsilon \psi_{k_x}(y). \quad (1.18)$$

¹¹We can ignore any terms proportional to $\mathbb{1}$ since they simply result in equal shifts of the energy for the valence and conduction band.

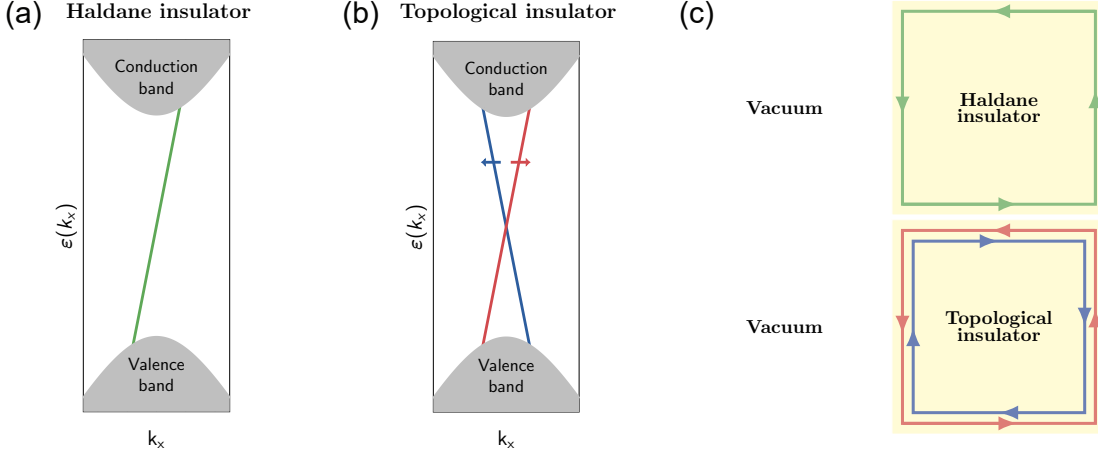


Fig. 1.4 **Edge states of the Haldane model and topological insulators:** a) The Haldane model possesses a single chiral edge mode with an approximately linear dispersion. For spinful fermions this mode is doubly degenerate. b) The model of Kane and Mele realises two time-reversal copies of the Haldane model, one for each spin polarisation in the z -direction. The helical edge modes have opposite Fermi-velocity for opposite spin polarisations $\pm s_z$. c) Edge modes in real space. Top: Since time-reversal symmetry is broken the edge mode of the Haldane insulator propagates in only one direction, it is topologically protected by the finite Chern number of the bulk. Bottom: In real space the chiral edge modes of Kane and Mele's model are two counter-propagating modes. Although the net Chern number is zero the edge states are protected by time reversal symmetry even when perturbations break s_z conservation (see next section)

Clearly the very large mass in the vacuum necessitates $\psi = 0$ in that region. On the other hand, within the bulk of the Haldane insulator $\psi_{k_x}(y)$ must satisfy Eq. 1.4 for $y > 0$ and $M(y) = M_H$. Assuming the mass is substantially larger than the kinetic energy $M_H \gg v_F k_x$ and requiring it decays at $y \rightarrow \infty$, we find $\psi_{k_x}(y)$ is well approximated by an eigenspinor of σ_x and the full unnormalised wave-function has the form

$$\psi_{k_x}(y) \approx \begin{pmatrix} 1 \\ \text{sign}(M_H) \end{pmatrix} e^{-ik_x x} e^{-\frac{|M_H|y}{v_F}}, \quad (1.19)$$

with energy $\varepsilon_{k_x} = \text{sign}(M_H)v_F k_x$. There are several interesting features of this mode: 1) It is localised to the edge, decaying exponentially into the bulk of the Haldane insulator – hence the name *edge mode*. 2) It is chiral with a pseudospin dependent on the sign of the mass M_H . 3) Similarly its dispersion is linear with a Fermi-velocity v_F that changes sign with the sign of the mass. The appearance of edge modes at the boundary between topological and non-topological states is a generic feature known as the *bulk-boundary correspondence*, which states that the number of left moving and right moving modes is equal to the change in Chern number at the boundary: $N_R - N_L = \Delta n$.

1.4 Topological insulators in 2d

In the last section we saw that the Haldane model provides an elegant mechanism that breaks time-reversal symmetry in graphene and at the same time retains the Bloch physics required to establish its Dirac nature. The result was an insulator with a non-zero Chern number and – of particular interest to our discussion of Dirac Hamiltonians – the establishment of an edge mode at the boundary between the Haldane-insulator and a trivial insulator. While a phenomenological model like Haldane’s is a nice sandbox for understanding the potential implications of topology, it is not easily realised in actuality (outside of artificial cold atom systems [39]). It therefore seems almost an impossibility that a material could achieve not one but two copies of the Haldane insulator and, in doing so, retain time-reversal symmetry so that the material is *intrinsically* topological. Amazingly that is exactly what Kane and Mele succeeded in finding [8]. Perhaps even more unexpected is that such a situation is not only possible but the base scenario for several materials [14].

The next two sections of this chapter deal with exactly these materials, known as *topological insulators*. Since in part, or in whole, chapters 2, 3, 4, and 5 will establish many theoretical and experimental properties of topological insulators, the present discussion is focussed on their realisation of edge or surface states with a Dirac-*like* dispersion.

1.4.1 The Kane-Mele Model

In our previous discussion of both graphene and the Haldane insulator we largely ignored the spin of the electron, since we were not interested in terms that couple to it. In contrast, Kane and Mele [8] considered a perturbation to the graphene Hamiltonian due to the relativistic spin-orbit (SO) interaction. In our valley-pseudospin basis above the SO coupling can be written $H_{\text{SO}} = \Delta_{\text{SO}} \tau_z \sigma_z s_z$, where we have introduced yet another Pauli matrix¹², s_z , the component of spin perpendicular to the lattice plane. To be precise, SO coupling enters into the graphene tight-binding Hamiltonian as second nearest neighbour couplings, similar to the way the Haldane model breaks time-reversal symmetry

$$H = H_{\text{G}} + it' \sum_{\mathbf{r}_A, \mathbf{i}} \left(c_{A,\alpha}^\dagger(\mathbf{r}_A + \boldsymbol{\nu}_i) s_{\alpha\beta}^z c_{A,\beta}(\mathbf{r}_A) + c_{B,\alpha}^\dagger(\mathbf{r}_B + \boldsymbol{\nu}_i) s_{\alpha\beta}^z c_{B,\beta}(\mathbf{r}_B) \right), \quad (1.20)$$

where summation over the indices α, β are implied.

We see that this perturbation is exactly the term that generated the mass in the Haldane model, with an extra matrix index from the spin s_z . As a result, the SO term respects both inversion and time reversal symmetry – since the $\tau_z \sigma_z$ and s_z are both odd under TR – but still results in an energy gap $\varepsilon_{\mathbf{k}} = \pm \sqrt{(v_F \mathbf{k})^2 + \Delta_{\text{SO}}^2}$. Quite incredibly we have found a realistic Hamiltonian that is the same as the Haldane insulator, with hopping phases $\phi = \frac{\pi}{2}$ for spin-up electrons and $-\frac{\pi}{2}$ for spin down electrons, thereby realising two

¹²Using \otimes for the Kronecker product becomes cumbersome for so many matrices, in what follows all products of Pauli-matrices imply the Kronecker product.

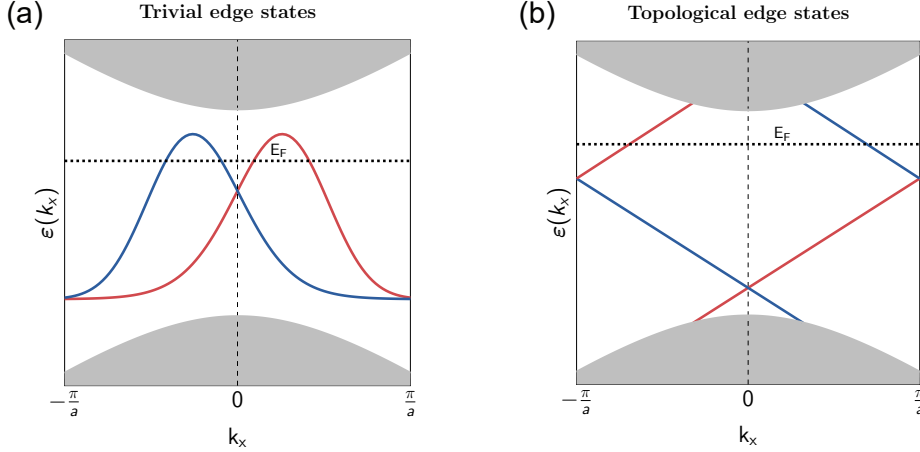


Fig. 1.5 **Trivial and topological edge states:** Due to Kramer's theorem all edge modes must be doubly degenerate at TRIMs – in 1d $k_x = \frac{\pi}{a}$ and $k_x = 0$ – and have a time reversal partner in the other half of the Brillouin zone $\pm k_x$. a) Trivial edge states are not protected by time-reversal symmetry and cross the Fermi-level an even number of times between the TRIMs at 0 and π/a , and so they can be pushed out of the bulk gap. b) Topological edge states are protected by time-reversal and cross the Fermi-level an odd number of times between the TRIMs [14].

copies of the Haldane Hamiltonian for each spin-species of electrons. Since the spin-orbit interaction of Kane and Mele's model respects all the symmetries of graphene it is not only a possible perturbation but *will* be present and the Dirac cones will be (slightly) gapped. Although we will see later in this section that several materials exist with high SO coupling and similar properties to Kane and Mele's model, luckily for its own applications as a 2d Dirac material the actual SO induced gap in graphene is extremely small ($\sim 42 \mu\text{eV}$ equivalent to $\sim 0.5 \text{ K}$ [45]).

1.4.2 \mathbb{Z}_2 topological insulators

Since Kane and Mele's model realises two copies of Haldane's model for each spin species, this also means that there are *two* edge states. At the edge between the insulator and the vacuum, where the mass vanishes, from the Haldane model we see these edge states are described by the 1d Dirac Hamiltonian

$$H = v_F s_z \sigma_x k_x, \quad (1.21)$$

which has eigenvalues $\varepsilon_k = \pm v_F k$, with opposite spins having the opposite sign Fermi-velocities (see Fig. 1.3). This spin momentum-locking of the *real* spin of the electron differentiates the Kane-Mele model and the Dirac Hamiltonian of topological insulator in general from, for instance, graphene where we saw the Pauli-matrices denoted a pseudospin resulting from the sublattice degree of freedom.

In the case of the Haldane insulator we saw that the edge state was guaranteed due to a finite Chern number. The two copies of Haldane's insulator in Kane and Mele's model have

1.4. Topological insulators in 2d

opposite Chern numbers n_\downarrow , n_\uparrow and so the system has a net zero Chern number $n_\uparrow + n_\downarrow = 0$, which had to be the case since time reversal is unbroken. It may therefore seem that any term coupling these two copies¹³ will destroy the edge modes. This is actually not the case and the edges are protected against such perturbations by time-reversal symmetry¹⁴ [9].

To understand why the edge modes remain even when a Rashba – or other s_z non-conserving term – is present we need to use Kramers' theorem which states that the time reversal operator satisfies $\Theta^2 = -1$ for a Hamiltonian of spin-1/2 electrons. Since our Hamiltonian is invariant under \mathcal{T} , if a non-degenerate eigenstate $|\psi\rangle$ existed then it must also be eigenstate of Θ , such that $\Theta|\psi\rangle = t|\psi\rangle$. Applying time-reversal again gives $\Theta^2|\psi\rangle = |t|^2|\psi\rangle = -|\psi\rangle$ which is impossible because it would require $|t|^2 = -1$. Therefore our original assumption that the eigenstate was non-degenerate had to have been wrong and all eigenstates of the Hamiltonian *must* be degenerate [15]. In particular any state at $+\mathbf{k}$ always has a time-reversal partner with opposite spin at $-\mathbf{k}$ – together known as a *Kramers pair* – unless time reversal symmetry is broken.

Kramers' theorem is obviously especially important for the points in the Brillouin zone which time reversal maps to themselves, known as time reversal invariant momenta (TRIM). At a TRIM Kramers' theorem enforces that the states are two-fold degenerate and two-bands must cross or touch. Away from these TRIMs spin-orbit coupling enables this degeneracy to be split. How the bands connect in-between TRIMs is extremely important for their stability: If the bands connect as a pair as in Fig. 1.5 then they can be pushed out of the gap by any perturbation obeying time-reversal symmetry. In contrast, if the bands disappear into the bulk conduction and valence bands as in Fig. 1.5, they cannot be removed [14].

From Fig. 1.5 we see that whether the edge state is robust or not can be diagnosed simply by counting the number of crossings of the Fermi-level in the positive ($q \geq 0$) quadrant of the Brillouin zone. If there are an even number of band crossings then the edge mode can be eliminated, if there is an odd number the edge is robust unless the perturbation is so large it closes the bulk band gap. For the 1d edge of our Kane-Mele model, above the TRIMs are $k_x = 0$ and $k_x = \pi/a$ and our edge mode crosses the Fermi-level in the positive quadrant only once. This means that any time-reversal preserving perturbation, even those that lead to a coupling of s_z such as the Rashba term above, does not destroy the edge mode [9].

The stability of the edge mode is a remarkable result: We started with two pairs of the Haldane model – a system that has a net zero Chern number – and showed that, even when they couple, the edge state is topologically protected by time reversal symmetry. Of course the integer valued Chern number – only non-zero when time reversal is broken – was not the topological invariant guaranteeing us our edge mode. We can also already

¹³An example of one such perturbation to the Kane-Mele model is a Rashba interaction $\lambda \hat{\mathbf{z}} \cdot (\mathbf{s} \times \mathbf{k})$ – which is allowed when $z \rightarrow -z$ mirror symmetry is broken by, for example, a substrate – this explicitly breaks s_z conservation and links the two copies of the Haldane model [9].

¹⁴We will see that this protection comes from the fact the \mathbb{Z} Chern number is broken down to a \mathbb{Z}_2 invariant $\nu = (n_\uparrow - n_\downarrow)/2 \pmod{2}$, with $n_{\uparrow,\downarrow}$ the individual Chern number of the up and down spins, respectively [16].

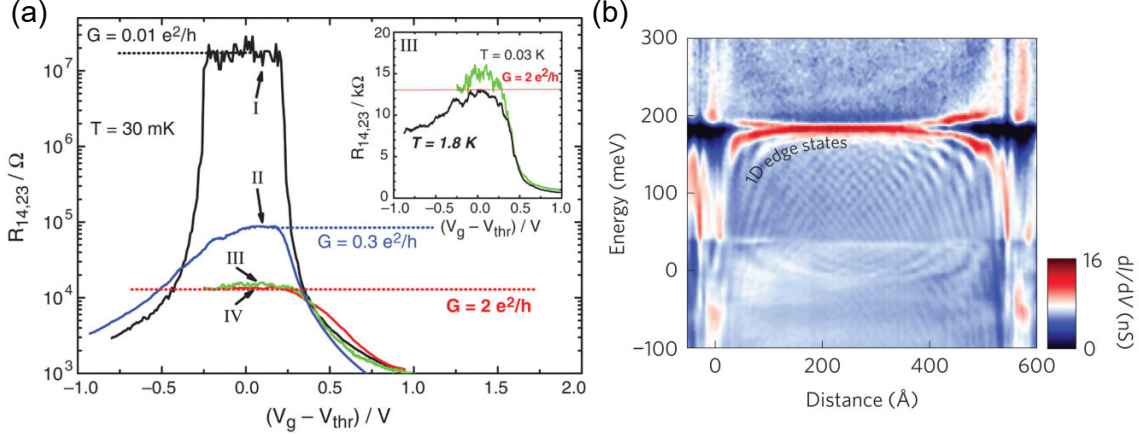


Fig. 1.6 **Topological insulators in 2d:** (a) Four terminal resistance measurements on HgTe quantum wells. In the bulk band gap and above a critical thickness $d_c \sim 6.3 \text{ nm}$ a single quantum of conductance $G = 2e^2/h$ is measured, characteristic of a clean 1d state as predicted by Bernevig, Hughes, and Zhang. (b) Scanning tunnelling spectroscopy measurements on a bismuth bilayer showing the presence of edge states (red) at the boundary of the bilayer and a bulk Bi-crystal. (a: Reproduced from König, Markus, et al. "Quantum spin Hall insulator state in HgTe quantum wells." *Science* 318.5851 (2007): 766-770.[17] b: Reproduced from Drozdov, Ilya K., et al. "One-dimensional topological edge states of bismuth bilayers." *Nature Physics* 10.9 (2014): 664. [47])

see from the fact the number of protected edge states is either 0 or 1, depending on the number Fermi-level crossings at the edge states, suggests that the topological invariant ν involved is \mathbb{Z}_2 in nature. Indeed similar to the Berry curvature we can define the matrix $w_{mn}(\mathbf{k}) = \langle u_m(\mathbf{k}) | \Theta | u_n(-\mathbf{k}) \rangle$ built from the eigenstates $|u_n(\mathbf{k})\rangle$ that are occupied and measures the effect of time-reversal on the Bloch states. At the TRIMs – which we label Γ_i – the fact that $\Theta^2 = -1$ guarantees that the “sewing matrix,” $w_{mn}(\mathbf{k})$ is anti-symmetric. Fu and Kane [46] showed that $\delta_i = \text{Pf}[w(\Gamma_i)]/\det(w(\Gamma_i)) = \pm 1$, where $\text{Pf}[\cdot]$ is the Pfaffian of the anti-symmetric matrix, correctly measures whether or not the edge band is protected. In 2d there are four TRIMs and we can define the topological invariant in terms of the product $(-1)^\nu = \prod_i \delta_i$, (for $i \in \text{TRIMS}$) [16]. We will see how this extends to 3d topological insulators in the next section.

1.4.3 Realisations of 2d TIs

As mentioned above, since carbon has a fairly low atomic number, its spin-orbit interaction is extremely weak and any topological insulator behaviour in graphene – or its heavier element analogs such as silicene – has still not been conclusively observed. Bernevig, Hughes, and Zhang (BHZ) [10] realised that heavier elements with much stronger spin-orbit coupling could achieve a similar band-inversion effect. In particular, bulk HgTe is a zero band-gap semiconductor with a zincblende structure and has an inversion between its p- and s- orbital bands. In contrast, CdTe is a trivial insulator with no band inversion but also realises a zincblende lattice with a slightly larger lattice constant.

BHZ's insight was that sandwiching a thin layer of HgTe between two slabs of CdTe the difference in lattice constants induces strain on the HgTe layer causing it to develop a gap. Above a critical thickness $d_c \sim 6.3$ nm the HgTe retains its band inversion of the p- and s- orbitals and so realises a 2d topological insulator. When the thickness is reduced below this critical value a phase transition occurs because the band inversion vanishes and the HgTe layer becomes a trivial insulator. These HgTe wells were very quickly synthesised by the Molenkamp group [17] who were able to show that, above the critical thickness, the conductance becomes quantised to $2e^2/h$ when the HgTe layer is tuned so that the Fermi-energy sits in the bulk band gap, indicating the presence of a 1d gapless edge state living within the bulk gap and capable of hosting a single quantum of conductance.

Since these first experiments HgTe has become well established as the archetypal 2d topological insulator. Another 2d topological insulator that has been experimentally realised is bilayer Bismuth Bi(111) where it was recently possible to directly observe the edge states by ARPES and STS revealing its Dirac nature [47] (see Fig. 1.6). Finally the layered material AlSb/InAs/GaSb/AlSb was also predicted [48] to have a small non-trivial bulk gap and subsequent experiments have found a $2e^2/h$ quantised conductance similar to HgTe quantum wells and consistent with the existence of topological edge states [49].

1.5 Topological insulators in 3d

The Haldane insulator was simply a type of *Chern insulator* meaning it is analogous to the quantum Hall states. Such states can only exist in two dimensions.¹⁵ It was quickly realised that it *was* possible to generalise the concept of a \mathbb{Z}_2 topological insulator to three dimensions [11–13]. Rather than the edge modes of their 2d cousins these materials host *surface states*. Just as in the case of a 2d topological insulator, these states appear at the boundary between the topological insulator and a trivial insulator, however these semi-metallic modes now cover the full two dimensional surface of the topological insulator. Unlike the 2d Dirac states of graphene, these surface states are governed by a Dirac equation where the spin entering the Dirac Hamiltonian is the *actual* spin of the electron and will require time-reversal to be broken in order to develop a gap. Again since several phenomena are discussed in later chapters, our current focus is on the properties and development of 3d TIs.

1.5.1 Topological invariant(s)

The surface states of a 3d crystal can be labelled by a 2d surface momentum $\mathbf{k} = (k_x, k_y)$. Just as in the case for 1d edge modes, at time-reversal invariant momenta, Kramers' theorem forces the surface states to be degenerate. In 2d there are four TRIMs located at $(0, 0)$, $(0, \pi/a)$, $(\pi/a, 0)$, and $(\pi/a, \pi/a)$ in the k_x - k_y plane. As in the case of edge modes, to ascertain the system's topology we need to find a way to describe the connectivity between these 4-points.

¹⁵Mathematically they are possible in *even* spatial dimensions.

Perhaps the most obvious way to create a 3d topological insulator is to stack layers of 2d TIs on top of each other [14]. The result is that the 1d Fermi-points of the edge states are expanded into a Fermi-lines that span the length of the Brillouin zone, with each branch having opposite spin. The system is therefore topological in one direction – along the direction of the stacked chiral edges – but not in the direction of the stacking. The surface state of stacked 2d TIs is therefore *not* protected against perturbations that couple different layers – for instance impurities that scatter isotropically in \mathbf{k} can cause localisation. For this reason these types of 3d TIs are known as *weak* topological insulators. To diagnose a weak topological insulator we can make use of the same topological invariant ν of the 2d TI on each surface orientation x, y, z and denote these weak indices by (ν_1, ν_2, ν_3) .

There is, however, one more degree of freedom ν_0 which is purely three dimensional. Similar to how ν in 2d is able to determine the number of 1d Fermi-crossings, ν_0 determines the number of 2d crossings of a Fermi-circle. To be precise, by constructing the same weaving matrix as in the 2d scenario, $w_{mn}(\mathbf{k}) = \langle u_m(\mathbf{k}) | \Theta | u_n(-\mathbf{k}) \rangle$, we can determine all 4 topological invariant's ν_i by once again considering $\delta(\Gamma_i) = \text{Pf}[w(\Gamma_i)]$, where Γ_i now labels the 8 TRIMs of the full 3d system. As expected $\nu_{i=1,2,3}$ is the same as considering the product of the 4 TRIMs of a 2d surface. For instance $(-1)^{\nu_3} = \delta(\Gamma_{0,0,\pi})\delta(\Gamma_{0,\pi,\pi})\delta(\Gamma_{\pi,0,\pi})\delta(\Gamma_{\pi,\pi,\pi})$ is the topological invariant associated with the plane perpendicular to z . The final topological invariant ν_0 is the product of $\delta(\Gamma_i)$ at *all* 8 TRIMs $(-1)^{\nu_0} = \prod_{i=1,\dots,8} \delta(\Gamma_i)$. Topological insulators in 3d are therefore able to be identified by their 4 topological invariants $(\nu_0; \nu_1, \nu_2, \nu_3)$.

The topological insulator with $\nu_0 = 1$ is special since it is topologically protected in all directions and, unlike a weak topological insulator, is stable against any time-reversal preserving perturbations. For this reason it is known as a *strong* topological insulator. In particular, its Fermi-surface is closed with a spin-texture that continuously rotates along a path around the TRIM where the surface bands cross (see Fig. 1.7d).

1.5.2 Surface state Hamiltonian

As was the case for 2d topological insulators, real 3d TI materials require strong spin-orbit coupling that can create a band inversion in the bulk. We already saw that, due to a strong SO coupling, Bismuth bilayers are 2d TIs that host 1d surface states. The vast majority of 3d TIs are also built upon Bismuth with the most prominent realisations being Bi_2Se_3 and Bi_2Te_3 which are (1;000) topological insulators (see Fig. 1.7). These materials have a complicated crystal structure but the important bulk bands close to the Fermi-energy can be expressed in terms of the four p_z orbitals on a the Bi and X the other element (e.g Se or Te) $|p_{z,\text{Bi}}^+, \uparrow\rangle$, $|p_{z,\text{Bi}}^-, \downarrow\rangle$, $|p_{z,\text{X}}^+, \uparrow\rangle$, and $|p_{z,\text{X}}^-, \downarrow\rangle$. In this basis, the general form of the bulk

1.5. Topological insulators in 3d

Hamiltonian in terms of the momentum operators \mathbf{p} is

$$H = \begin{pmatrix} M_{\mathbf{k}} & A_1 p_z & 0 & A_2 p_- \\ A_1 p_z & -M_{\mathbf{k}} & A_2 p_- & 0 \\ 0 & A_2 p_+ & M_{\mathbf{k}} & -A_1 p_z \\ A_2 p_+ & 0 & -A_1 p_z & -M_{\mathbf{k}} \end{pmatrix} = \tau_z \mathbb{1} M_{\mathbf{k}} + \tau_x \boldsymbol{\sigma} \cdot \mathbf{p}, \quad (1.22)$$

where $p_{\pm} = p_x \pm ip_y$ [50]. For simplicity, in what follows we consider an isotropic version so that $A_1 = A_2 = A$ and assume the mass $M_{\mathbf{k}}$ to be momentum independent.

We repeat a similar procedure as in 2d TIs to look for the 3d TI's surface states. First we consider the half-volume $z > 0$ with a boundary between the TI and the vacuum at $z = 0$ and so the wave-function vanishes in $z < 0$ due to the large excitation gap of the vacuum. Since translation symmetry is maintained in the x and y directions, k_x and k_y remain good quantum numbers. This also informs our choice of ansatz wavefunction as plane waves in the x - and y - directions and the surface state decaying into the bulk exponentially in the z -direction, so that

$$\psi = e^{-\frac{Mz}{A}} e^{i(k_x x + k_y y)} \begin{pmatrix} \phi \\ \xi \end{pmatrix}, \quad (1.23)$$

where ϕ and ξ are spinors corresponding to the electron spin. Similar to the 2d case, after inserting this into the Hamiltonian (1.22), we find a relation between the spinors ϕ and ξ requiring $M\phi = iM\sigma_z\xi$ for the mass term to vanish.

Using this relation we can reduce the Hamiltonian to a 2×2 matrix equation of the surface state

$$\begin{aligned} H_{\text{surf}}^z &= v_F(k_x \sigma_x + k_y \sigma_y) i \sigma_z \\ &= v_F(k_x \sigma_y - k_y \sigma_x), \end{aligned} \quad (1.24)$$

where we defined $v_F = A\hbar$. In the same way that the edge state of a 2d TI realised a 1d Dirac Hamiltonian with a single Dirac cone centered at $k = 0$ in the edge Brillouin zone, the surface state of a 3d TI realises a 2d Dirac Hamiltonian with a single Dirac cone located at $\mathbf{k} = 0$ in the surface Brillouin zone. Once again the Pauli-matrices $\boldsymbol{\sigma}$ entering the Dirac Hamiltonian indicate the component of the *real* spin now within the x - y plane – or the plane which holds the surface.¹⁶ The constant energy contours (CEC) of Eq. (1.24) are circular and the spin of the electron is locked *perpendicular* to the momentum \mathbf{k} (see Fig. 1.7d). We will see the important effect that 2d *spin-momentum locking* has on the response of the surface state to disorder and in transport in the coming chapters.

Although Eq. (1.24) is an excellent approximation to the surface state Hamiltonian close to the Dirac point, the full Hamiltonian of a bismuth based topological insulator contains further terms at higher order in the momentum $|\mathbf{k}|$. There are three major sources

¹⁶In chapter 6 we will discuss how the surface states change when we confine them on a wire.

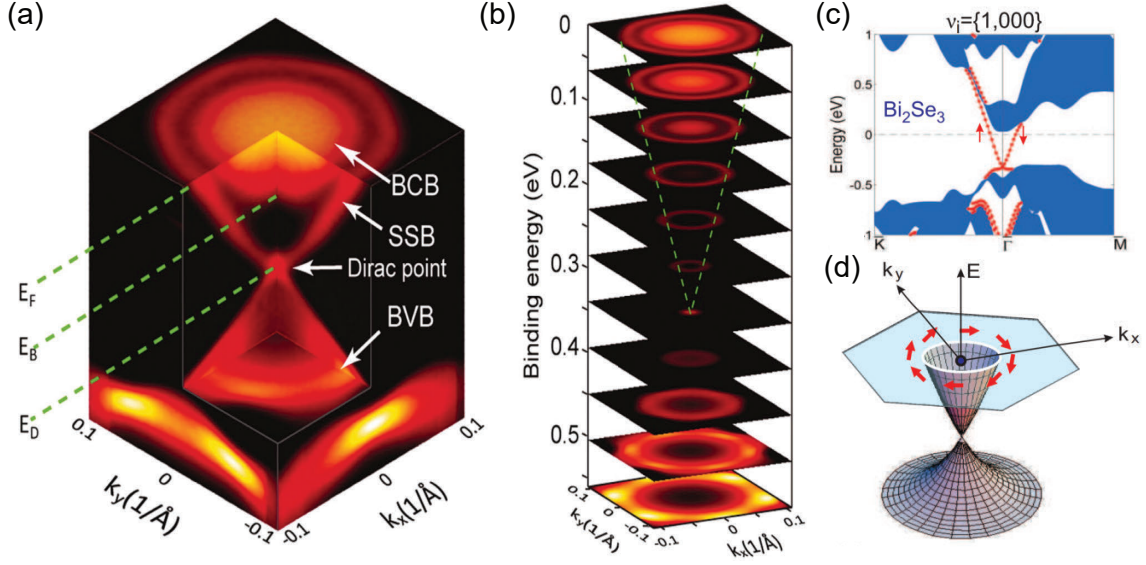


Fig. 1.7 **Surface states of Bi_2Se_3** : (a-b) Band structure of undoped Bi_2Se_3 measured by angle-resolved photo-emission spectroscopy (ARPES). (a) The bulk conduction band (BCB), surface state band (SSB), and bulk valence band (BVB) are labelled. (b) CEC of the band structure for increasing energies showing the enlargement of the surface state Fermi-ring from a node at the Dirac point to a hexagon close to the BCB. c) Surface electronic structure of Bi_2Se_3 calculated using the local density approximation. Red lines indicate surface states, blue lines are bulk states. d) Schematic of the spin polarized surface state dispersion in Bi_2X_3 TIs (a-b: Reproduced from “Massive Dirac Fermion on the Surface of a Magnetically Doped Topological Insulator.” Y. L. Chen, et al. Science 329, 659 (2010). [53] c-d: Reproduced from “Colloquium: topological insulators.” Hasan, M. Zahid, and Charles L. Kane. Reviews of modern physics 82.4 (2010): 3045. [14])

of higher order terms [51]: 1) Close to the Dirac point the Hamiltonian is approximately particle-hole symmetric, however, as the energy moves away from the Dirac point the bulk bands break this particle hole symmetry causing a quadratic term $E(\mathbf{k}) = \frac{|\mathbf{k}|^2}{2m^*} \mathbb{1}$, which is diagonal in spin-space. 2) The breaking of particle-hole symmetry also causes a curvature of the surface bands, which can be included in Eq. (1.24) by an effective \mathbf{k} dependence of the Fermi-velocity $v_F(\mathbf{k}) = v_F(1 + \beta|\mathbf{k}|^2)$. 3) The surface structures of both Bi_2Se_3 and Bi_2Te_3 have 6 fold rotational symmetry, this means that any higher order terms in $|\mathbf{k}|$ must respect this hexagonal symmetry and have the form $\delta H_{k^3} = \frac{\lambda}{2}(k_+^3 + k_-^3)$, where $k_{\pm} = k_x \pm ik_y$. This latter effect is visible in Fig. 1.7b and is especially relevant in Bi_2Te_3 [52] (see Fig. 1.8).

1.5.3 Realisations of 3d TIs

Despite the concept of topological insulators stemming from 2d TIs and in particular the Kane-Mele model on graphene, many more 3d TIs have been found than 2d TIs [16]. The first 3D TI was $\text{Bi}_{1-x}\text{Sb}_x$, a strong (1;111) topological insulator. Experiments by Hsieh *et al.* [18] found that a $\sim 10\%$ Sb doping resulted in surface states with a total of 5 band crossings (see Fig. 1.8), indicating the topological nature of the surface states.

1.5. Topological insulators in 3d

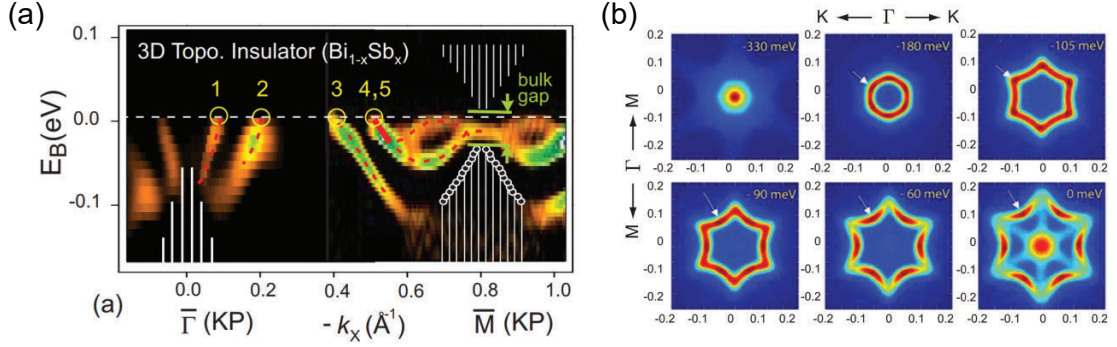


Fig. 1.8 **Surface states of 3d TIs:** a) Surface states of the first 3d TI $\text{Bi}_{1-x}\text{Sb}_x$, the surface band crosses the Fermi-level 5 times indicating the topological nature of the surface state. b) The surface states of Bi_2Te_3 measured by scanning tunnelling microscopy, showing a significant hexagonal warping at higher energies. (a: Reproduced from "A topological Dirac insulator in a quantum spin Hall phase." Nature 452.7190 (2008): 970. Hsieh, David, et al. [18] b: Reproduced from "STM imaging of electronic waves on the surface of Bi_2Te_3 : topologically protected surface states and hexagonal warping effects." PRL 104.1 (2010): 016401. Alpichshev, Zhanybek, et al. [52])

Unfortunately such a large number of surface modes also have the drawback that there are additional states at the Fermi-energy. As can be seen in Fig.1.8a the bulk gap in $\text{Bi}_{1-x}\text{Sb}_x$ is also extremely small (~ 25 meV [18]). Both the small gap and the complexity of the surface states make experimental investigations of $\text{Bi}_{1-x}\text{Sb}_x$ difficult.

As mentioned above Bi_2Se_3 and Bi_2Te_3 are certainly the most prominent and best examples of 3d TIs [14]. In particular the almost ideal surface Dirac cones of Bi_2Se_3 produce excellent ARPES images due to its chemical purity [16, 53], as shown in Fig. 1.7. In addition Bi_2Se_3 has a large bulk band gap of ~ 0.3 eV allowing a wide window of energies where the physics is surface dominated [16]. Bi_2Te_3 has a slightly smaller band gap of ~ 0.17 eV. As mentioned at the end of the previous section, its main drawback is a significant distortion of the surface state due to hexagonal warping [52], as shown in Fig. 1.8b.

Despite its near perfect Dirac cone, Bi_2Se_3 is actually naturally electron doped and so has a large bulk conductivity. This can be worked around by mixing various dopings of Te, Se, and/or Sb to reduce the materials doping and produce bulk insulating samples. The best *insulating* 3d TI is $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$, which will make many more appearances in later chapters [54].

A smorgasbord of other 3d TI materials, not directly relevant to this thesis, have been synthesised and predicted. For instance, it is thought that the ever controversial SmB_6 is an exotic Kondo topological insulator [55]. A detailed review of many TI materials in 2d and 3d can be found in Ref. [16].

1.6 Dirac and Weyl semi-metals in 3d

So far we have only discussed scenarios where 1d and 2d Dirac Hamiltonians appear in condensed matter systems. We have seen that symmetries play an important role in protecting these crossings. For instance, in graphene a combination of time-reversal and inversion symmetry meant that a mass could only develop if one of these symmetries was broken. In topological insulators time-reversal by itself provided the equivalent protection. Perhaps unsurprisingly, band crossings are easier to achieve in 3 dimensions due to the increased number of tuning parameters. As early as 1929 von Neumann and Wigner [56] recognised that it is possible to write the general Hamiltonian, up to constants of energy, of any two band system as

$$H = \begin{pmatrix} H_{11} & H_{12}^* \\ H_{12} & -H_{11} \end{pmatrix} = \text{Re}\{H_{12}\}\sigma_x + \text{Im}\{H_{12}\}\sigma_y + H_{11}\sigma_z, \quad (1.25)$$

with $H_{11} \in \mathbb{R}$ and $H_{12} \in \mathbb{C}$. The energy of such a system is $E = \pm\sqrt{H_{11}^2 + |H_{12}|^2}$. Hence in order to achieve a band crossing the off-diagonals – representing couplings between the two bands – must vanish at the same time the diagonal is equal to zero. In total this means that a band crossing can only occur when three constraints are satisfied $\text{Re}H_{12} = \text{Im}H_{12} = H_{11} = 0$, which is possible to achieve “accidentally” when there are 3 or more parameters – e.g. momenta – to adjust, as occurs in 3 dimensions (or higher). On the other hand, band crossings only exist in lower than 3 dimensions if further symmetries enforce a crossing, as we saw in graphene and topological insulators.

1.6.1 Dirac semi-metals

Of course the original high-energy Dirac equation contains a 3 spatial-dimension Hamiltonian, which we saw in Eq. (1.6) is chiral in the massless limit. The high energy Dirac equation, however, is a 4×4 matrix Hamiltonian and so describes a 4 band system, containing two copies of the two band model considered by Wigner and von Neumann in Eq. (1.25). In the massless limit this matrix Hamiltonian splits into blocks of chirality such that

$$H_D = \begin{pmatrix} 0 & -\boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & 0 \end{pmatrix}. \quad (1.26)$$

The eigenstates of this Hamiltonian are doubly degenerate with eigenvalues $E_{\pm} = \pm|\mathbf{p}|^2$ and opposite chiralities. Any term on the diagonal of the Hamiltonian will couple the two chiralities and result in an energy gap, for the Dirac equation this occurs for a finite mass and the “gap” is equal to the mass-energy mc^2 .

A lattice version of this four-band model occurs in, for instance, the material $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ which forms a cubic lattice with a bipartite rock salt structure with a Pb or Sn atom on one site of the unit cell and Te on the other [26, 27]. Its Brillouin zone is therefore that of the FCC lattice and it has minimal band gaps at the L-points of this Brillouin

1.6. Dirac and Weyl semi-metals in 3d

zone (more information can be found in chapter 8). For our present discussion what is important is that each L-point's low energy physics can be described in a 4 component basis $\{|\psi_{L,1}^+\rangle, |\psi_{L,1}^-\rangle, |\psi_{L,2}^+\rangle, |\psi_{L,2}^-\rangle\}$, where \pm is the sublattice index and 1, 2 indicate the total angular momentum $J_z = \pm \frac{\hbar}{2}$ along the direction between the Γ -point at the center of the Brillouin zone and the L-point. Due to the crystalline symmetries of $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$, within this basis, the low energy Hamiltonian expanded about a specific L-point has the form [27]

$$H = \hbar \begin{pmatrix} m & 0 & -iv'k_z & -v(ik_x + k_y) \\ 0 & m & v(ik_x - k_y) & -iv'k_z \\ iv'k_z & -v(ik_x - k_y) & -m & 0 \\ v(ik_x + k_y) & iv'k_z & 0 & -m \end{pmatrix}, \quad (1.27)$$

with m depending on the relative doping of Pb and Sn. The eigen-energies of this Hamiltonian are $E_{\pm} = \pm \hbar \sqrt{v^2 |\mathbf{k}|^2 + m^2}$ and for a critical doping $x \approx 0.17$ the mass m disappears [27]. In the massless limit we see that each 2×2 block on the off-diagonals of the Hamiltonian satisfy the von Neumann-Wigner constraints from Eq. (1.25) at $\mathbf{k} = 0$ and the four bands cross at a point. Such a material, with crossing of 4 bands at a single point, is known as a **3d Dirac semi-metal**. The node that occurs at the crossing of these 4 bands in a 3d system is known as a **Dirac point**.

1.6.2 Weyl semi-metals

The Hamiltonian of a Dirac semi-metal involves 4 parameters and so, in a similar manner to the 2d Dirac Hamiltonians we have already seen, can only occur when enforced by additional symmetries that prevent the coupling of its two 2×2 Hamiltonians that individually satisfy the von Neumann-Wigner conditions. Obviously the easiest way to satisfy these conditions with just two bands would be to look for materials with just a single copy of the blocks on the diagonal of the high-energy Dirac Hamiltonian, such as

$$H_W^{(+)} = +\hbar v(\boldsymbol{\sigma} \cdot \mathbf{k}). \quad (1.28)$$

This Hamiltonian describes what is known as a **3d Weyl semi-metal** and by construction realises a two band crossing with a node at $\mathbf{k} = 0$, known as a **Weyl point**. The eigenfunctions in the vicinity of this Weyl point are spinors with an angular dependence [57]

$$\psi(\theta, \varphi) = \begin{pmatrix} \sin\left(\frac{\theta_{\mathbf{k}}}{2}\right) \\ -\cos\left(\frac{\theta_{\mathbf{k}}}{2}\right) e^{i\varphi_{\mathbf{k}}} \end{pmatrix}, \quad (1.29)$$

where θ is the azimuthal angle and φ the polar angle. This is the same spinor that describes a spin-1/2 fermion in a magnetic field, which we know results in an Aharnov-Bohm phase. To see what this means for our Weyl point consider the Berry connection $\mathbf{A}(\mathbf{k}) = i \langle \psi_{\mathbf{k}} | \nabla_{\mathbf{k}} | \psi_{\mathbf{k}} \rangle = \cos^2(\theta/2) \hat{\boldsymbol{\varphi}}$. Transforming to Cartesian coordinates and taking

the curl we find that the Berry curvature is $\mathbf{\Omega}_k = \hat{\mathbf{k}}/2k^2$ [57]. Like we saw in 2d, to find the Berry phase of a single node we need to integrate this over a sphere around the Weyl point, the result is a phase of 2π . The natural interpretation is that, just as the flux from a magnetic field causes an Aharnov-Bohm phase, the Weyl node acts as a source of Berry flux i.e. a monopole of “magnetic field”, which here is the Berry connection. If instead we had chosen the opposite sign in front of our two band Hamiltonian – i.e. $H_W^{(-)} = -\hbar v(\boldsymbol{\sigma} \cdot \mathbf{k})$ – we would then find that the Berry phase was -2π , signalling a monopole with the opposite charge.

The Brillouin zone is a 3-torus which means that the only way that two bands can cross at one point is if they re-cross at another point or else the bands wouldn’t be periodic.¹⁷ As a result Weyl nodes must appear in *pairs* and these pairs must have opposite chirality. The net charge of the chirality of all Weyl nodes in the Brillouin zone must therefore add to zero. That Weyl points always come in pairs is known as the “fermion doubling theorem” [58]. When a magnetic field is applied, this high energy connection between pairs of Weyl nodes leads to what is known as the chiral anomaly [23] (we will discuss this and other magnetoresistance phenomena in 3d Dirac systems in chapter 7).

The doubling theorem also has consequences for the stability of Weyl nodes. If we could eliminate an individual Weyl point, then its opposite chirality partner would cause there to be a net charge over the whole Brillouin zone, which is not allowed. Therefore the only way to eliminate a Weyl point is to adjust parameters such that two points of opposite chirality are brought together and their charges cancel. Weyl nodes therefore have a quite different protection from mass formation than the symmetry enforced crossings that we have seen before. Since it is not protected by symmetry this is known as an “accidental degeneracy” and is unique to 3 dimensional materials [23].

We have argued that Weyl points always come in pairs. Let’s now consider what happens when other symmetries are present in the system. For instance if we assume that a Weyl point appears in the Brillouin zone at a location \mathbf{k}_0 then time reversal symmetry requires $H_W^{(\pm)}(\mathbf{k} + \mathbf{k}_0) = H_W^{(\pm)\dagger}(-\mathbf{k} - \mathbf{k}_0)$ so that another Weyl point with the *same* chirality must appear at $-\mathbf{k}_0$. Since Weyl points always come in pairs of opposite chirality, there must be at least 4 Weyl points in the system if time reversal is present. On the other hand inversion symmetry requires $H_W^{(\pm)}(\mathbf{k} + \mathbf{k}_0) = H_W^{(\pm)}(-\mathbf{k} - \mathbf{k}_0)$, this means a Weyl point of *opposite* chirality must appear at $-\mathbf{k}_0$ if inversion symmetry is present. If *both* inversion and time reversal symmetry are present then it is clear from the above discussion that a pair of opposite chirality Weyl points must be located at \mathbf{k}_0 and a pair at $-\mathbf{k}_0$, in other words Weyl points separated in momentum space *cannot* exist unless either time-reversal or inversion symmetry is broken and if both are present the only allowed band crossings are Dirac points [2].

¹⁷Of course bands could also cross multiple times, as long as it is in total an even number of crossings.

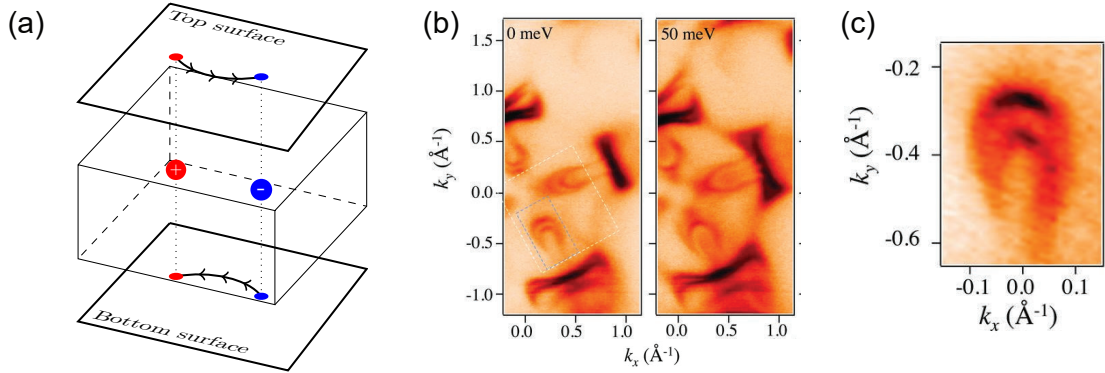


Fig. 1.9 **Surface states of Weyl semi-metals:** a) Schematic of the Fermi-arcs on the surface of a Weyl semi-metal. The arcs connect the projection of the Weyl nodes onto the top and bottom surface in the opposite sense. b-c) Fermi-arcs of TaAs measured by ARPES. The chiral charge of the Weyl points in TaAs is ± 2 leading to two Fermi-arcs on its surface. (b-c: Reproduced from "Discovery of a Weyl fermion semimetal and topological Fermi arcs " Science 349.6248 (2015): 613-617. Xu, Su-Yang, et al. [25])

1.6.3 Surface states

In the case of the Haldane model and topological insulators we have seen that a non-trivial bulk leads to edge modes in 1d and surface modes in 2d. Weyl semi-metals also have surface states,¹⁸ in contrast to these previous examples, the surface states of Weyl semi-metals are 1d Fermi-arcs i.e. the surface states are 2 dimensions lower than their 3d bulk [2]. To see why, consider a system with Weyl points at $\pm \mathbf{k}_0$ which are opposite monopoles of Berry flux. If we now consider the impact of a surface by taking a cut in a momentum-space – for ease consider a plane perpendicular to \mathbf{k}_0 – the location of this plane in the 3d Brillouin zone will determine its topology; in either the region of the Brillouin zone between $+\mathbf{k}_0$ to $-\mathbf{k}_0$ or its complement region between $-\mathbf{k}_0$ and \mathbf{k}_0 there is a net Berry flux and a non-trivial topology, in the other region there is no net flux and therefore a trivial topology. In combination the 2d “bulks” of these cuts describe a 1d edge between trivial and non-trivial regions, with the edge appearing at the projection of one Weyl node onto the surface and disappearing at the other. On the opposite surface this edge mode connects the two points in the opposite direction. These Fermi-arcs have been observed experimentally in TaAs¹⁹ [25], as shown in Fig. 1.9.

1.7 Other Dirac systems

To end this chapter we briefly discuss other system where Dirac-like excitations appear. In particular it is important to note that the Dirac equation can describe the low energy excitations of a variety of systems which are not necessarily electronic in nature.

¹⁸Surface states can also occur in so-called *topological* Dirac semi-metals where the two copies of Weyl points at a momentum \mathbf{k}_0 have their topological partners located at $-\mathbf{k}_0$. This results in a pair of Fermi-arcs between the two Dirac points [59].

¹⁹Unfortunately both TaAs and the 3d Dirac semi-metal Cd_3As_2 – arguably two most prominent 3d Dirac/Weyl semi-metals – require crystal growers to handle arsenic, which is highly toxic.

1.7.1 Topological crystalline insulators

We discussed in detail how time-reversal symmetry protects the surface states of \mathbb{Z}_2 topological insulators such as Bi_2Se_3 . It is, however, possible to obtain the same protection via crystalline symmetries. For instance if a material has a mirror symmetry M then the Bloch states can be chosen such that they are also eigenstates of M with mirror eigenvalues $\pm i$. The total Chern number can then be represented as the sum of the Chern number of each eigenvalue $N = N_+ + N_-$. Even if $N = 0$, however, it is possible for the *mirror Chern number* $N_M = \frac{N_+ - N_-}{2}$ to be a non-zero integer.

In fact we have already seen a material which has a non-zero mirror Chern number when we considered the Hamiltonian of $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ in Eq. (1.27). In $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ the mass m undergoes an inversion as a function of doping x , indicating that either PbTe or SnTe is topologically non-trivial [26]. The number of band inversions and corresponding number of surface bands at the Fermi-level is, however, even and so the mass inversion does not signal the \mathbb{Z}_2 topology of a standard TI. Rather it is found that SnTe has a non-zero mirror Chern number, indicating it is topologically non-trivial. On the other hand PbTe has mirror Chern number 0 and is therefore trivial. As in the case of \mathbb{Z}_2 TIs, the surface states of topological crystalline insulators – which can only exist on mirror-symmetric surfaces – are well described by a 2d Dirac Hamiltonian, with each surface containing an even number of Dirac cones. (More information can be found in [27]).

1.7.2 d-wave superconductors

Over the last 30 years there has been significant theoretical and experimental effort to understand the high temperature superconductors (HTSC), such as the cuprates. One of the defining features of these HTSCs is an order parameter $\Delta_{\mathbf{k}} = \Delta_0(\cos(k_x a) - \cos(k_y a))$, which has a $d_{x^2-y^2}$ symmetry rather than the s-wave symmetry of most superconductors that can be described by a standard BCS theory.

A lot of evidence now points to the cuprates being well described as 2d layers governed by a Hamiltonian $H_d = \varepsilon_{\mathbf{k}}\tau_z + \Delta_{\mathbf{k}}\tau_x$, where $\varepsilon_{\mathbf{k}}$ is the electronic dispersion in the normal state [2]. Here the matrix structure of the Hamiltonian is defined in terms of the Nambu spinor $\Psi = (\psi_{\uparrow}, \psi_{\downarrow}^{\dagger})$ and have a total energy $E_{\mathbf{k}} = \pm\sqrt{\varepsilon_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}$. In most cuprates the locations of zeros of the dispersion $\varepsilon_{\mathbf{k}}$ correspond to the location of zeros of the order parameter $\Delta_{\mathbf{k}}$. If we expand about these nodal points in $E_{\mathbf{k}}$ it should be clear that the Bogliubov quasi-particles obey a Dirac dispersion and that the d-wave superconductors can also be included in the Dirac materials class (more information can be found in [2, 27, 60]).

1.7.3 Majorana metals

Another example of emergent Dirac physics occurs in the low-energy descriptions of certain magnetic systems. Certainly the most famous of these is the Kitaev model [32]. The Kitaev model exists on a honeycomb lattice and induces frustration, a necessary requirement for spin liquids behaviour, via exchange couplings with an easy-axis dependent on bond

direction. The nearest neighbour spins follow the Hamiltonian $H_{\text{Kit.}} = \sum_{\gamma} K_{\gamma} S_i^{\gamma} S_j^{\gamma}$, where each $\gamma = x, y, z$ also indexes one of the 3 bond directions in the honeycomb lattice.

Surprisingly this Hamiltonian can be exactly solved by a mapping of each spin in terms of four *Majorana* fermion operators $S_i^{\gamma} = i a_i^{\gamma} c_j$ [30]. After performing this calculation, the Hamiltonian becomes a nearest neighbour tight binding Hamiltonian of Majorana's hopping on the honeycomb lattice. As such these Majorana's on the honeycomb lattice form a "metal" which inherits the same Dirac physics as electrons in graphene. In particular their low energy excitations are determined by a Dirac equation expanded about the K-points [30]. More recently it has been shown that other tri-coordinated lattices in 3d can be solved in a similar manner and, depending on the lattice, are capable of hosting Weyl and Dirac points, as well as other exotic band structures such as nodal lines [31]. Interestingly, unlike for complex fermions, due to the peculiar nature of their majorana Weyl points, it is possible to have a Weyl semi-metal phase in these materials when both time-reversal and inversion symmetry are present [61]. (More information can be found in [32]).

Summary and outlook

In this chapter we have seen that the Dirac equation appears in many condensed matter settings. The variety of systems that achieve a Dirac Hamiltonian distinguish them as a separate class of materials with a unique phenomenology. Many of these materials are some of the most interesting in modern condensed matter physics, such as graphene and topological insulators. The presence of a Dirac dispersion also results in many novel consequences, such as the spin-momentum locking of the surface states of topological insulators and the unique Fermi-arcs on the surface of Weyl semi-metals. The appearance of Dirac fermions in condensed matter is also deeply interconnected with topology and symmetries, which protect the Dirac dispersions from developing a mass.

Now that we have an overview for the remainder of this thesis we switch to a more detailed discussion of the consequences of disorder in Dirac systems. With a particular focus on 3d topological insulators and Dirac semi-metals and their transport properties.

Disorder in Dirac materials

In the previous chapter we found that the Dirac equation governs the low energy excitations in a wide variety of materials where symmetries and topology enforce the crossing of two bands close to the Fermi-energy. In this chapter we consider the impact of disorder in those systems. Disorder is an unavoidable fact of life in real materials¹ and can stem from many sources, for instance impurities – examples include foreign or charged atoms –, or lattice dislocations – such as surface step edges –, both break the translational symmetry of the system. The main impact of these defects is to act as scattering centers for electrons, altering their trajectories as they travel through a material. Ultimately it is disorder that governs the electronic properties of most metallic states at low temperatures, such as the transport properties we discuss in chapter 5.

There is, however, another more theoretical reason to consider disorder: The physics of disordered systems is often significantly *more* exciting and collective phenomena become *more* important. For example, Anderson localisation means that the electrons' wavefunctions in a one or two dimensional electron gas are localised by the presence of even a minuscule amount of disorder [62]. The integer quantum Hall effect, which was able to measure so accurately the von Klitzing constant $\frac{e^2}{h}$ that it is now a basis of the SI measurement system [63], actually *requires* a finite amount of disorder to be visible since impurities have the effect of both extending the Hall plateaus as a function of magnetic field and localising the bulk states [64]. This means that disorder is not some undesirable feature of real world systems but a key component that is often of acute theoretical importance.

In order to have a concrete setting to discuss disorder effects we will consider the 2d surface states of a 3d topological insulator. Although much of our discussion can be generalised to other materials such as graphene [65] and to three dimensional materials which we will discuss in chapters 6 and 7. One of the main impacts of the Dirac physics is its protection from backscattering which we discuss in the first section of this chapter and introduce for the first time the T-matrix formalism that will be used throughout this thesis.² We will then discuss the impact of a finite density of impurities and the concept of disorder averaging. Having developed the theoretical tool kit to treat disorder, in the latter parts of this chapter we will calculate the self-energy and density of states of a disordered surface of a 3d TI.

¹Some materials such as the such as the 2d TI HgTe quantum wells in Fig. 1.6 can actually be incredibly clean [17] but these are the exception rather than the rule. This is also related to the reason why they exhibit a quantised conductance.

²We will follow a similar format to Ref. [66] but applied to Dirac states rather than an electron gas.

2.1 Disordered 3d topological insulator surface

The main impact of an impurity on the surface of a topological insulator surface is to break translation symmetry. In practice what this means is that an impurity enables the plane-waves that form the eigenstates of a clean TI surface to elastically scatter from a momentum \mathbf{k} to a new momentum \mathbf{k}' . To demonstrate this we start with the Hamiltonian of a clean surface state close to the Dirac point, as described by Eq. (1.24) in the previous chapter. To this surface we add a single spin-less impurity which acts as a local potential $\langle \mathbf{r} | \hat{U} | \mathbf{r} \rangle \equiv U(\mathbf{r}) = u(\mathbf{r} - \mathbf{r}_i) \mathbb{1}$, for the moment we consider only a single impurity and therefore without loss of generality we can take the impurity to be located at the origin $\mathbf{r}_i = 0$. We will discuss the impact of a finite density of impurities later.

It is normally easier to perform calculations in momentum-space since the eigenstates of the clean system can be labelled by their momentum \mathbf{k} . On the surface of a topological insulator it is important to remember that the clean eigenstates also contain a spinor due to spin-momentum locking which in the conduction (+) and valence (−) bands can be expressed,

$$\langle \mathbf{r} | \mathbf{k}, + \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} +ie^{-i\theta_{\mathbf{k}}} \\ 1 \end{pmatrix} e^{i\mathbf{k} \cdot \mathbf{r}} \quad \text{and} \quad \langle \mathbf{r} | \mathbf{k}, - \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -ie^{-i\theta_{\mathbf{k}}} \\ 1 \end{pmatrix} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (2.1)$$

where $\theta_{\mathbf{k}}$ is the polar angle in the $k_x - k_y$ plane with $\theta_{\mathbf{k}} = 0$ corresponding to $\mathbf{k} \parallel \hat{\mathbf{k}}_x$ and the conduction and valence bands have left-handed and right-handed chiralities, respectively.

It should already be clear that the spinor structure of these eigenstates makes them very different to those of a 2DEG. In particular if we consider the matrix element that appears in Fermi's golden rule of the impurity for a scattering from \mathbf{k} to \mathbf{k}' in a single event we find

$$|\langle \mathbf{k}, \pm | \hat{U} | \mathbf{k}', \pm \rangle|^2 = |u_{\mathbf{k}-\mathbf{k}'}|^2 \cos^2 \left(\frac{\theta_{\mathbf{k}} - \theta_{\mathbf{k}'}}{2} \right), \quad (2.2)$$

where $u_{\mathbf{k}-\mathbf{k}'}$ is the Fourier transform of the real-space impurity potential. Unlike a 2DEG, where such a matrix element is isotropic and proportional only to $|u_{\mathbf{k}-\mathbf{k}'}|^2$, the spin-momentum locking of the surface state causes an additional angular dependence. In particular back-scattering $\mathbf{k} \rightarrow -\mathbf{k}$ is completely forbidden and scatterings close to 180° are strongly suppressed compared with those that preserve the electron's trajectory. This is a unique feature of Dirac fermions which are said to be protected from back-scattering.

In general we will consider two different types of impurity potentials: Point-like impurities, $u_{\mathbf{q}}^\delta = u_0$ and screened charged impurities $u_{\mathbf{q}}^c = \frac{u_0}{q^2 + \kappa_0^2}$, where κ_0 is the Thomas-Fermi screening vector and u_0 is the strength of the impurity. The former does not have any \mathbf{k} dependence and only the $\theta_{\mathbf{k}}$ angular dependence due to spin-momentum locking enters the matrix element, the latter strongly prefers small momenta scatterings and will further reduce the overlap between states on opposite sides of the Fermi-surface.

2.1. Disordered 3d topological insulator surface

Let us now consider what were to happen when we break time-reversal symmetry and generate a mass term $m\sigma_z$ in the surface state Hamiltonian such that Eq. (1.24) becomes

$$H_{\text{Surf}}^m = v_F(k_x\sigma_y - k_y\sigma_x) + m\sigma_z. \quad (2.3)$$

As expected there is now a gap of size $2m$ and the eigenvalues are $\varepsilon_{\mathbf{k},m}^\pm = \pm \sqrt{v_F^2(k_x^2 + k_y^2) + m^2}$. More importantly the eigenstate's spinor is no longer an eigenspinor of some combination of the σ_x and σ_y Pauli-matrices. Instead the spin now cants out of the surface plane parallel to $\hat{\mathbf{z}}$. This means the surface state is no longer fully topological protected and

$$|\langle \mathbf{k}, \pm | \hat{U} | \mathbf{k}', \pm \rangle|^2 = |u_{\mathbf{k}-\mathbf{k}'}|^2 \left(\frac{m^2}{\varepsilon_{\mathbf{k},m}^2} + \frac{v_F k^2 \cos\left(\frac{\theta_{\mathbf{k}} - \theta_{\mathbf{k}'}}{2}\right)^2}{\varepsilon_{\mathbf{k},m}^2} \right), \quad (2.4)$$

in particular there is a finite back-scattering rate $\xi = |\langle \mathbf{k}, \pm | \hat{U} | -\mathbf{k}, \pm \rangle|^2 = \frac{m^2}{\varepsilon_{\mathbf{k},m}^2}$. We also see, however, that for large energies $\varepsilon_{\mathbf{k},m} \approx v_F |\mathbf{k}| \gg m$ the protection is approximately restored. This of course means that graphene, where we showed SO coupling leads to a small but finite gap in the Dirac spectrum, is still strongly protected from back-scattering unless the Fermi-energy is very close to the Dirac point (assuming impurities approximately respect the symmetries of graphene's Hamiltonian [65]).

2.1.1 The T-matrix

We have seen that a single scattering event described by the matrix elements like $\langle \mathbf{k} | \hat{U} | \mathbf{k}' \rangle$ already tells us a lot of information about our Dirac fermions in the presence of disorder. This matrix elements, however, only give us the first order in a perturbative series of increasing orders of scattering from the disorder potential \hat{U} [66]. To find the full effect of the disorder on the eigenstates $|\psi\rangle$ we need to solve the eigenenergy equation of the clean Hamiltonian plus the impurity

$$(\hat{H}_{\text{surf}} + \hat{U}) |\psi\rangle = \varepsilon |\psi\rangle. \quad (2.5)$$

In real space the eigenstates of the full Hamiltonian must satisfy the integral equation

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \int d\mathbf{r}_1 \mathbf{G}_0(\mathbf{r} - \mathbf{r}_1) \mathbf{U}(\mathbf{r}_1) \psi(\mathbf{r}_1), \quad (2.6)$$

where $\psi(\mathbf{r})$ are 2-spinors and the (retarded) real space Green's function $\mathbf{G}_0(\mathbf{r} - \mathbf{r}')$ is the Fourier transform of

$$\mathbf{G}_0(\mathbf{k}, \omega) = (\omega - H_{\text{surf}} + i\delta)^{-1} = \frac{\omega + v_F(k_x\sigma_y - k_y\sigma_x)}{\omega^2 - v_F^2 k^2 + i\delta}, \quad (2.7)$$

which is a 2×2 matrix in the Pauli basis of the TI surface. The matrix structure of the Green's function encodes the spin-momentum locking of the surface states.³ It should be clear that this integral equation can be expanded perturbatively in orders of $\mathbf{U}(\mathbf{r})$ by repeated insertions of the full eigenstate $\psi(\mathbf{r})$, leading to the series [66]

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \int d\mathbf{r}_1 \mathbf{G}_0(\mathbf{r} - \mathbf{r}_1) \mathbf{U}(\mathbf{r}_1) \psi_0(\mathbf{r}_1) \quad (2.8)$$

$$+ \int d\mathbf{r}_1 d\mathbf{r}_2 \mathbf{G}_0(\mathbf{r} - \mathbf{r}_1) \mathbf{U}(\mathbf{r}_1) \mathbf{G}_0(\mathbf{r}_1 - \mathbf{r}_2) \mathbf{U}(\mathbf{r}_2) \psi_0(\mathbf{r}_2) + \dots \quad (2.9)$$

$$= \psi_0(\mathbf{r}) + \int d\mathbf{r}_1 d\mathbf{r}_2 \mathbf{G}_0(\mathbf{r} - \mathbf{r}_1) \mathbf{T}(\mathbf{r}_1, \mathbf{r}_2) \psi_0(\mathbf{r}_2), \quad (2.10)$$

where $\mathbf{T}(\mathbf{r}_1, \mathbf{r}_2)$ is known as the **T-matrix**. The T-matrix – also a 2×2 matrix – is the sum of all terms representing repeated scattering events of the clean system with the impurity $\mathbf{U}(\mathbf{r})$. For calculations it is normally easier to switch to momentum space where the T-matrix can be expressed in terms of itself and the momentum-space Green's function

$$\begin{aligned} \mathbf{T}_{\mathbf{k}, \mathbf{k}'}(\omega) &= u_{\mathbf{k}-\mathbf{k}'} + \sum_{\mathbf{k}_1} u_{\mathbf{k}-\mathbf{k}_1} \mathbf{G}^0(\mathbf{k}_1) u_{\mathbf{k}_1-\mathbf{k}'} \\ &\quad + \sum_{\mathbf{k}_1, \mathbf{k}_2} u_{\mathbf{k}-\mathbf{k}_1} \mathbf{G}^0(\mathbf{k}_1) u_{\mathbf{k}_1-\mathbf{k}_2} \mathbf{G}^0(\mathbf{k}_2) u_{\mathbf{k}_2-\mathbf{k}'} + \dots \\ &= u_q + \sum_{\mathbf{k}_1} u_{\mathbf{k}-\mathbf{k}_1} \mathbf{G}^0(\mathbf{k}_1) \mathbf{T}_{\mathbf{k}_1, \mathbf{k}'}. \end{aligned} \quad (2.11)$$

We can express the T-matrix diagrammatically as

$$T_{\mathbf{k}, \mathbf{k}'}(\omega) = \text{[Diagram: vertical dashed line with diamond]} = \text{[Diagram: vertical dashed line with X]} + \text{[Diagram: triangle with dashed/solid lines]} + \text{[Diagram: triangle with two internal dashed lines]} + \dots, \quad (2.12)$$

where the dashed lines represent scattering events and the solid lines Green's functions. This diagrammatic representation will be an extremely useful tool in later chapters.

The first order of the T-matrix, which gave us our golden rule matrix elements, corresponds to *weak scattering*, meaning that it is valid for low impurity strengths u_0 and can be easily calculated for any impurity potential u_q . For point-like impurities the T-matrix is independent of momentum and so the full series can be summed and Eq. (2.11) simplifies to

$$\mathbf{T}^\delta(\omega) = u_0 + u_0 \sum_{\mathbf{q}} \mathbf{G}^0(\mathbf{q}, \omega) \mathbf{T}^\delta(\omega). \quad (2.13)$$

³The Green's function and the T-matrix both depend on energy ω , however for clarity of the equations this index is implied. We can do this because impurity scattering is elastic and so the energy does not change after each scattering event.

2.2. Impurity averaging

Taking the continuum limit we find that the T-matrix for point-scatterers is

$$\begin{aligned} \mathbf{T}^\delta(\omega) &= u_0 \left(\omega \mathbb{1} - u_0 \int \frac{d\mathbf{q}}{(2\pi)^2} G^0(\mathbf{q}, \omega) \right)^{-1} = u_0 \omega \mathbb{1} \left(1 - \frac{u_0}{2\pi} \int_0^\Lambda dk \frac{k}{\omega^2 - v_F^2 k^2} \right)^{-1} \\ &= u_0 \mathbb{1} \left(1 - \frac{u_0 \omega}{4\pi v_F^2} \ln \left[\frac{-\Lambda^2}{\omega^2} \right] \right)^{-1}, \end{aligned} \quad (2.14)$$

where we introduced the ultra-violet cut-off⁴ $\Lambda \gg \omega$ and used the fact that the angular integral over the off-diagonals of the Green's function is zero [65]. This last fact ensures that the full T-matrix of a diagonal impurity is *also* diagonal in spin-space and so the matrix element $\langle \mathbf{k}, \pm | \mathbf{T}(\omega) | -\mathbf{k}, \pm \rangle = 0$ and back-scattering remains forbidden, even due to multiple scattering events from the same impurity. At first sight this might seem surprising since an initial state with momentum \mathbf{k} that scatters to an intermediate state at \mathbf{k}' is not prohibited from then scattering to the state at $-\mathbf{k}$, however, this path is cancelled by an equivalent path starting at \mathbf{k} scattering to $-\mathbf{k}'$ and ending at $-\mathbf{k}$. The net result after considering all paths, at all orders in the scattering potential, is that the protection from back-scattering is retained.⁵

The momentum independence of point-like scatterers is, of course, a special case and for a general potential $u_{\mathbf{q}}$ the full T-matrix has to be calculated numerically, especially given the added complication of the Green's function's matrix structure. Nonetheless, even for a momentum dependent scatterer, the above argument always holds and for every possible back-scattering path there is always an equivalent destructively interfering path, unless time-reversal symmetry is broken either by an external field or by the impurities themselves – we will see an example of the latter situation in chapter 5.

2.2 Impurity averaging

So far we have only considered a single impurity, of course on the surface of any real topological insulator there will be a finite density of impurities $\frac{N_{\text{imp}}}{A} = n_{\text{imp}}$. Unless the system has been prepared otherwise, it is reasonable to assume that the disorder is entirely randomly distributed and that the full disorder potential is $\mathbf{U}(\mathbf{r}) = \sum_{\mathbf{r}_i} u(\mathbf{r} - \mathbf{r}_i) \mathbb{1}$, where the position of the impurity \mathbf{r}_i is chosen at random from a constant distribution and $i \in \{1, \dots, N_{\text{imp}}\}$ so that they are completely uncorrelated. We now extend the solution for scattering of the wave-function from a single impurity, Eq. (2.6), to this finite density case. Equivalent to considering the full wave-function $\psi(\mathbf{r})$ we can (and will) consider $\mathbf{G}(\mathbf{r}_1, \mathbf{r}_2)$, the full Green's function defined by

$$\mathbf{G}(\mathbf{r}, \mathbf{r}') = \mathbf{G}_0(\mathbf{r} - \mathbf{r}') + \int d\mathbf{r}_1 \mathbf{G}_0(\mathbf{r} - \mathbf{r}_1) \mathbf{U}(\mathbf{r}_1) \mathbf{G}(\mathbf{r}_1, \mathbf{r}'). \quad (2.15)$$

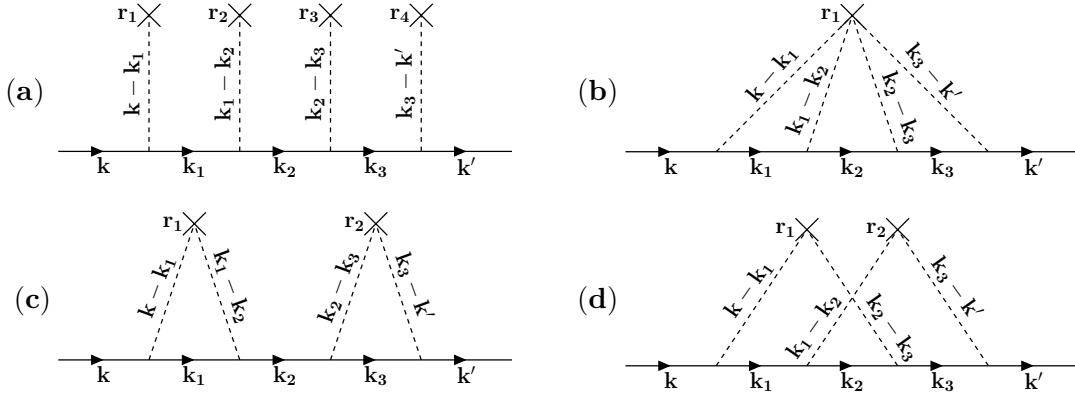
⁴Physically such a cut-off is just a reminder that our Dirac dispersion is only a low-energy description.

⁵This protection from back-scattering at all orders ensures that the wave-functions can't localise and we can have a genuine two dimensional electronic state [9].

Performing the same expansion as in the single impurity case, now by repeated insertion of the free Green's function, we obtain a perturbative expansion of the full Green's function with the n th order term given by [67]

$$\mathbf{G}^{(n)}(\mathbf{r}, \mathbf{r}') = \sum_{i_1, \dots, i_n}^{N_{\text{imp}}} \int d\mathbf{r}_1 \dots \mathbf{r}_n \mathbf{G}_0(\mathbf{r} - \mathbf{r}_1) u(\mathbf{r}_1 - \mathbf{r}_{i_1}) \mathbf{G}_0(\mathbf{r}_1 - \mathbf{r}_2) \dots u(\mathbf{r}_n - \mathbf{r}_{i_n}) \mathbf{G}_0(\mathbf{r}_n - \mathbf{r}'), \quad (2.16)$$

which has a contribution from all possible combinations of scatterings from n events. For instance the electron could scatter from n completely different impurities – corresponding to $\mathbf{r}_{i_1}, \dots, \mathbf{r}_{i_n}$ different for every scattering event ((a) below) – or n times from the same impurity – corresponding to $\mathbf{r}_{i_1} = \dots = \mathbf{r}_{i_n}$ being the same for all events ((b) below). For example some of the 4th order terms – labelling the Green's function by momentum (see below) – are:



Given the huge numbers of impurities in macroscopic sized samples and the impossibility of knowing all of their locations, the expansion of the Green's function Eq. (2.16) is clearly not a particularly useful object if one wants to do any actual calculations.

To make progress in solving Eq. (2.16) it is important to recognise that the large numbers of impurities and the size of the system involved actually work to our advantage. Kohn and Luttinger realised that a macroscopic sample can be split into subsystems of a size set by the decoherence length l_ϕ , which is governed by all scattering processes in the system and can be interpreted as the distance an electron needs to travel before its phase is completely randomised. For large systems, here surfaces $A \gg l_\phi^2$, there are still a macroscopic number of subsystems, such that throughout the sample virtually every possible impurity configuration within a subsystem is realised [67].

For macroscopic samples, we can therefore (normally) consider the wave-function as being *averaged* over all possible impurity distributions, denoted by $\langle \cdot \rangle_{\text{imp}}$. When the distribution of impurities is uniform this simply means performing the integral $\frac{1}{A} \int d\mathbf{r}_i$ for each impurity position \mathbf{r}_i . The advantage of this is that translational symmetry is *on average* restored to the system. In particular the Green's function after impurity averaging is diagonal in momentum \mathbf{k} , such that $\langle \mathbf{G}(\mathbf{k}, \mathbf{k}') \rangle_{\text{imp}} \equiv \mathcal{G}(\mathbf{k}) \sim \delta_{\mathbf{k}, \mathbf{k}'}$. This does *not*, however, mean that the effects of the impurities are forgotten.

2.2. Impurity averaging

2.2.1 Kohn-Luttinger expansion

Let us now consider what happens to the n -th order expansion of our Green's function in Eq. (2.16) when we impurity average [68]. In particular we want to consider the contributions to $\mathcal{G}(\mathbf{k})$ after Fourier transforming. Each scattering event, \mathbf{r}_{ij} , contributes to this Fourier transform a phase $e^{i\mathbf{r}_{ij} \cdot \mathbf{k}_i}$. The impurity average over these phases force the total incoming and out-going momenta to each impurity to sum to zero, such that

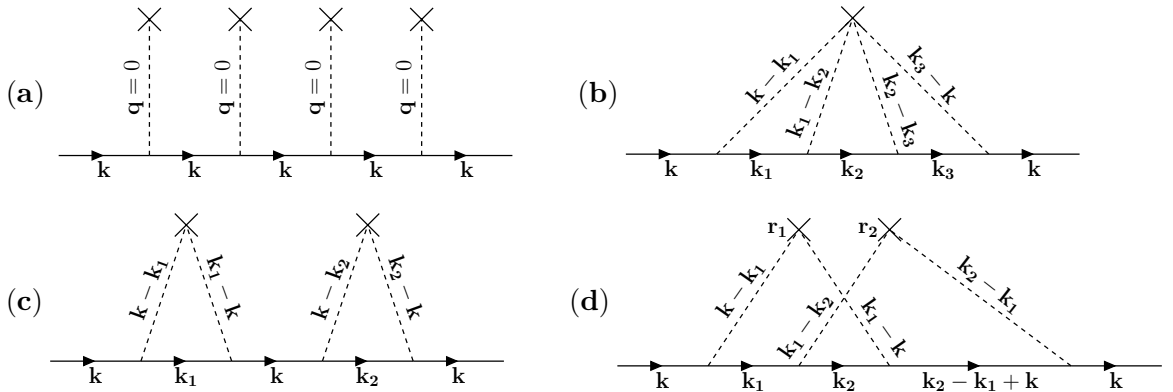
$$\left\langle \sum_j e^{i\mathbf{r}_{ij} \cdot \mathbf{k}_i} \right\rangle_{\text{imp}} = \frac{N_{\text{imp}}}{A} \delta_{\sum_i^n \mathbf{k}_i=0} + \frac{N_{\text{imp}}(N_{\text{imp}}-1)}{A^2} \delta_{\sum_i^m \mathbf{k}_i=0} \delta_{\sum_m^n \mathbf{k}_i=0} + \mathcal{O}(n_{\text{imp}}^3), \quad (2.17)$$

where due to the large numbers of impurities we can approximate $\frac{N_{\text{imp}}(N_{\text{imp}}-1)}{A^2} \approx n_{\text{imp}}^2$ and similarly for higher order terms in N_{imp} . This simply says that diagrams where impurity lines cross as diagram (d) above, have more constraints on the allowed \mathbf{k}_i and are therefore at a higher order in impurity density than the if all scattering events were from a single impurity, as for example diagram (b).

The end result is that the n -th order impurity averaged Green's function can be expanded in impurity density

$$\begin{aligned} \mathcal{G}(\mathbf{k}) = n_{\text{imp}} \mathbf{G}^0(\mathbf{k}) & \left(u_{q=0} + \int \frac{d\mathbf{k}_1}{(2\pi)^2} |u_{\mathbf{k}-\mathbf{k}_1}|^2 \mathbf{G}^0(\mathbf{k}_1) \right. \\ & \left. + \int \frac{d\mathbf{k}_1}{(2\pi)^2} \frac{d\mathbf{k}_2}{(2\pi)^2} u_{\mathbf{k}-\mathbf{k}_1} u_{\mathbf{k}_1-\mathbf{k}_2} u_{\mathbf{k}_2-\mathbf{k}} \mathbf{G}^0(\mathbf{k}_1) \mathbf{G}^0(\mathbf{k}_2) + \dots \right) \mathbf{G}^0(\mathbf{k}) \\ & + n_{\text{imp}}^2 \mathbf{G}^0(\mathbf{k}) \left(u_{q=0} \mathbf{G}^0(\mathbf{k}) u_{q=0} + u_{q=0} \int \frac{d\mathbf{k}_1}{(2\pi)^2} |u_{\mathbf{k}-\mathbf{k}_1}|^2 \mathbf{G}^0(\mathbf{k}_1) + \dots \right) \mathbf{G}^0(\mathbf{k}) \\ & + \mathcal{O}(n_{\text{imp}}^3). \end{aligned} \quad (2.18)$$

Luckily this is much more easily understood diagrammatically where we just need to ensure the incoming and out-going momenta at all vertices sum to zero. The diagrams above (a-d) after impurity averaging are:



2.3 The self-energy

The key terms in the Kohn-Luttinger expansion, Eq. (2.18), are the *irreducible diagrams*. These are all diagrams where a Green's function's line cannot be cut in half such as in (b) and (d), but not (a) and (c). These diagrams allow us to write down the Dyson equation for the impurity averaged Green's function – diagrammatically a double fermion line – as

$$\mathcal{G}(\mathbf{k}) = \mathbf{G}^0(\mathbf{k}) + \mathbf{G}^0(\mathbf{k})\Sigma_{\mathbf{k}}(\omega)\mathcal{G}(\mathbf{k}) = \text{---}\mathbf{k}\text{---} + \text{---}\mathbf{k}\text{---}\bigcirc\Sigma\text{---}\mathbf{k}\text{---}, \quad (2.19)$$

where we have defined the self-energy $\Sigma_{\mathbf{k}}(\omega)$, which is the sum of all irreducible diagrams with the two external $\mathbf{G}^0(\mathbf{k})$ removed

$$\Sigma_{\mathbf{k}}(\omega) = n_{\text{imp}} \left(\begin{array}{c} \text{---}\mathbf{k}\text{---} \\ \times \\ \text{---}\mathbf{k}\text{---} \\ q=0 \end{array} + \begin{array}{c} \text{---}\mathbf{k}-\mathbf{k}_1\text{---} \\ \times \\ \text{---}\mathbf{k}_1-\mathbf{k}\text{---} \\ \text{---}\mathbf{k}_1\text{---} \end{array} + \begin{array}{c} \text{---}\mathbf{k}-\mathbf{k}_1\text{---} \\ \times \\ \text{---}\mathbf{k}_1-\mathbf{k}_2\text{---} \\ \text{---}\mathbf{k}_2-\mathbf{k}\text{---} \\ \text{---}\mathbf{k}_1\text{---} \end{array} + \dots \right) \quad (2.20)$$

$$+ n_{\text{imp}}^2 \left(\begin{array}{c} \text{---}\mathbf{k}-\mathbf{k}_1\text{---} \\ \times \\ \text{---}\mathbf{k}_1-\mathbf{k}_2\text{---} \\ \text{---}\mathbf{k}_2-\mathbf{k}\text{---} \\ \text{---}\mathbf{k}_1\text{---} \end{array} \begin{array}{c} \text{---}\mathbf{k}_1-\mathbf{k}\text{---} \\ \times \\ \text{---}\mathbf{k}_2-\mathbf{k}_1\text{---} \\ \text{---}\mathbf{k}_2-\mathbf{k}_1+\mathbf{k}\text{---} \\ \text{---}\mathbf{k}_1\text{---} \end{array} + \dots \right) + \mathcal{O}(n_{\text{imp}}^3).$$

In the self-energy we recognise the lowest order term in the expansion of impurity density is simply $n_{\text{imp}}\mathbf{T}_{\mathbf{k},\mathbf{k}}$, the T-matrix as defined in Eq. (2.11). This lowest order in density approximation is known as the **non-crossing approximation** [66], since at these order there are no diagrams with crossed impurity lines from different scattering centers.

Rearranging the Dyson equation, Eq. (2.19), to solve for the full disorder averaged Green's function we obtain

$$\mathcal{G}(\mathbf{k}) = (\mathbf{G}^0(\mathbf{k})^{-1} - \Sigma_{\mathbf{k}}(\omega))^{-1}, \quad (2.21)$$

where we take the matrix inverse. The self-energy is therefore a measure of the broadening of the eigenstates in comparison to the clean system. In fact we can assign the **scattering time** $\frac{1}{\tau_{\mathbf{k}}} = -2\text{Tr}\{\text{Im } \Sigma_{\mathbf{k}}(\omega)\}$ which is the average time between scattering events. Physically the reason why $\Sigma_{\mathbf{k}}(\omega) = n_{\text{imp}}\mathbf{T}_{\mathbf{k},\mathbf{k}}(\omega)$ at lowest order in impurity density is now understandable: When the impurities are dilute, the eigenstates are in the vicinity of only one impurity at a time and scatter from it, potentially multiple times, before moving on and being influenced by other impurities. We can also make use of the scattering time $\tau_{\mathbf{k}}$ to define the impurity **mean free path** $l_{\mathbf{k}} = v_F\tau_{\mathbf{k}}$.

It's worth taking a moment to understand exactly what a “dilute” density of impurities n_{imp} means in this context. In the Kohn-Luttinger expansion Eq. (2.17) we saw that higher orders in impurity density n_{imp} lead to more constraints on the momenta \mathbf{k}_i . At a given

2.3. The self-energy

Fermi-energy the typical momentum is given by the Fermi-momentum k_F , as a result each additional constraint on momentum reduces the available scattering phase space by $\sim 1/k_F$. On the other hand the typical length scale for scattering is the mean free path at the Fermi-energy l_{k_F} , which is the same as saying the size of the disc in phase space provided by an impurity has a size governed by $1/l_{k_F}$. As such there is a competition between the increase in the size of the scattering phase-space a new impurity provides and the decrease in phase space caused by the additional constraints on momentum after impurity averaging, with the ratio given by $\sim 1/(k_F l_{k_F})$. In general $1/k_F$ is comparable to a lattice spacing⁶ and the mean free path is normally several 10s of lattice spacings, therefore the ratio $1/(k_F l_{k_F}) \ll 1$ and we can safely neglect higher orders in n_{imp} [69].

2.3.1 Self-consistent T-matrix

We have seen that the non-crossing approximation for the self-energy includes all diagrams involving only a single scatterer and so is proportional to the T-matrix as given in Eq. (2.11). We can include many more diagrams by returning to our definition of the T-matrix⁷ and defining it self-consistently, such that

$$\frac{\Sigma(\omega)}{n_{\text{imp}}} \approx \mathbf{T}^\delta(\omega) = u_0 + u_0 \sum_{\mathbf{q}} \mathcal{G}(\mathbf{q}) \mathbf{T}^\delta(\omega) = \text{---}\times\text{---} + \text{---}\blacklozenge\text{---}, \quad (2.22)$$

where we have replaced the free Green's function $G^0(\mathbf{k})$ with the full impurity averaged Green's function $\mathcal{G}(\mathbf{k})$, which itself depends on the T-matrix via the self-energy. This is known as the **self-consistent T-matrix** approximation.

As discussed above, the T-matrix is diagonal in spin-space at all orders and, as a result, so is the self-energy $\Sigma(\omega) = \Sigma(\omega)\mathbb{1}$. Therefore we can obtain the self-consistent equation for the self-energy by re-calculating T-matrix, as in Eq. (2.14), with $\omega \rightarrow \omega - \Sigma(\omega)$ [65], leading to

$$\Sigma(\omega) = n_{\text{imp}} u_0 \left(1 - \frac{u_0}{4\pi v_F^2} [\omega - \Sigma(\omega)] \ln \left[\frac{-\Lambda^2}{(\omega - \Sigma(\omega))^2} \right] \right)^{-1}. \quad (2.23)$$

In general the self-energy in this self-consistent equation must be calculated numerically. There are, however, two important limits: Weak and strong scattering, which we now discuss.

⁶This is not true when we are very close to the Dirac point since there $k_F \rightarrow 0$. Practically, however, it is valid except for a very small region of chemical potential close to the Dirac point.

⁷For the rest of this chapter we will consider point-like scatterers but the discussion can be easily extended for momentum dependent potentials.

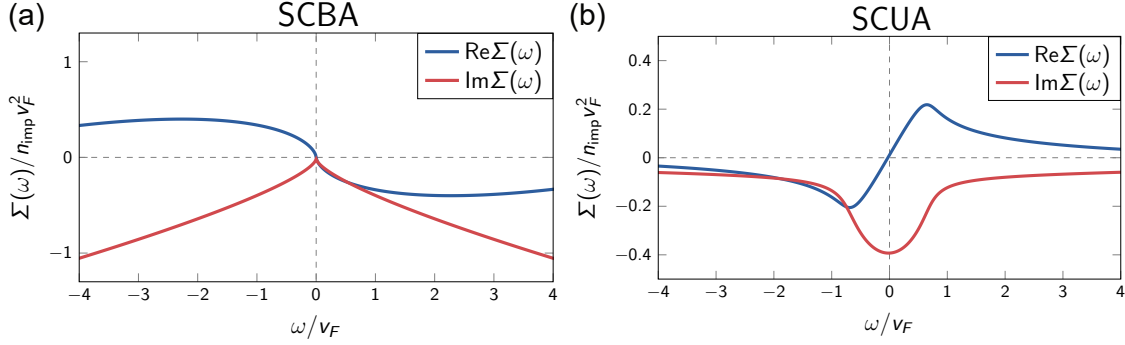


Fig. 2.1 **Self-energy in strong and weak scattering limits:** The real and imaginary parts of the self-energy for point-like scatterers in the self-consistent Born and unitary limits. (Parameters: $n_{\text{imp}} = 1/4\pi$, $\Lambda = 10$, and $v_F = 1$)

2.3.2 Self-consistent Born approximation

We see from the full Green's function, Eq. (2.21), that the first term in the T-matrix expansion $n_{\text{imp}} u_0$ can be incorporated into a shift of the energy ω . The next term defines the **self-consistent Born approximation** (SCBA) for weak scattering. When u_0 is small the self-energy reduces to

$$\Sigma_{\text{SCBA}}(\omega) \approx n_{\text{imp}} u_0^2 \sum_{k_1} \mathcal{G}(\mathbf{q}) = n_{\text{imp}} \quad , \quad (2.24)$$

which is the same as Eq. (2.22) but we have replaced the full T-matrix with the bare potential u_0 .

Switching to the continuum limit and performing the momentum integral we find

$$\Sigma(\omega) \approx \frac{n_{\text{imp}} u_0^2}{4\pi v_F^2} [\omega - \Sigma(\omega)] \ln \left[\frac{-\Lambda^2}{(\omega - \Sigma(\omega))^2} \right]. \quad (2.25)$$

As in [65] in the low energy regime, $\omega \ll \Lambda e^{-2\pi v_F^2/n_{\text{imp}}}$, we can use a small ω expansion which gives $\Sigma(\omega) \approx -i\Lambda e^{-2\pi v_F^2/n_{\text{imp}}}$. This means that the self-energy is finite (but very small), even at the Dirac point⁸ which we will see in the next section also means the density of states at the Dirac point is finite.

⁸As stated above, the non-crossing approximation breaks down at the Dirac point and crossed diagrams can become important. For weak scattering crossed diagrams in the self-energy are anyway of a higher order. In particular for calculations of quantities like the conductivity and the density of states such diagrams only have the effect of renormalising the energy scale Γ_0 [65]. In general the NCA is actually able to capture the qualitative physics – but not quantitative – of the a Dirac system coupled to impurities in the metallic regime far from the Dirac point as well as in the impurity dominated regime close to the Dirac point [65].

2.3.3 Self-consistent Unitary approximation

The opposite limit is the limit of *strong* scattering and is known as the **self-consistent unitary approximation** (SCUA) [65]. Here we can neglect unity in the self-consistent T-matrix equation, Eq. (2.26) giving

$$\Sigma(\omega) = -n_{\text{imp}} \left(\frac{1}{4\pi v_F^2} [\omega - \Sigma(\omega)] \ln \left[\frac{-\Lambda^2}{(\omega - \Sigma(\omega))^2} \right] \right)^{-1}. \quad (2.26)$$

The real and imaginary parts of both the self-consistent unitary and self-consistent Born approximation are plotted in Fig. 2.26.

2.4 Density of states

Now that we have the machinery to deal with disorder on our topological insulator surface lets use it to calculate the density of states. This is simply the imaginary part of the full Green's function summed over all of momentum space

$$\rho(\omega) = -\frac{1}{\pi} \text{Tr} \left\{ \text{Im} \int \frac{d\mathbf{k}}{(2\pi)^2} \mathcal{G}(\mathbf{k}) \right\}, \quad (2.27)$$

where we trace over spin-space since we want all states, regardless of spin orientation, at the energy ω . This is the same integral in Eq. (2.14), using that result, we see that the density of states is

$$\rho(\omega) = \frac{1}{2\pi^2 v_F^2} \text{Im} \left\{ [\omega - \Sigma(\omega)] \ln \left[\frac{-\Lambda^2}{(\omega - \Sigma(\omega))^2} \right] \right\}. \quad (2.28)$$

If there is no self-energy, corresponding to the clean system, the imaginary part of the logarithm provides a factor $\text{sgn}(\omega)i\pi$ such that the clean density of states is $\rho_0(\omega) = \frac{|\omega|}{2\pi v_F^2}$. In contrast, the density of states when disorder is present is plotted for the the two approximations above in Fig. 2.2. We see that in both cases the density of states tracks that of the clean system and there is a finite density of states at the Dirac point. In particular the density of states in the SCUA has an impurity dominated regime close to the Dirac point, where the impurities create a large number of new states [70]. We will discuss these further in chapter 5.

Summary and outlook

In this chapter we have seen how the presence of impurities alters a topological insulator's surface. In particular we introduced the T-matrix, which represents multiple scattering from an impurity. We also introduced the concept of disorder averaging, this will be important in enabling us to obtain the conductivity, which we will discuss in chapter 4. As a pedagogical example we used this T-matrix to calculate the disorder averaged self-energy

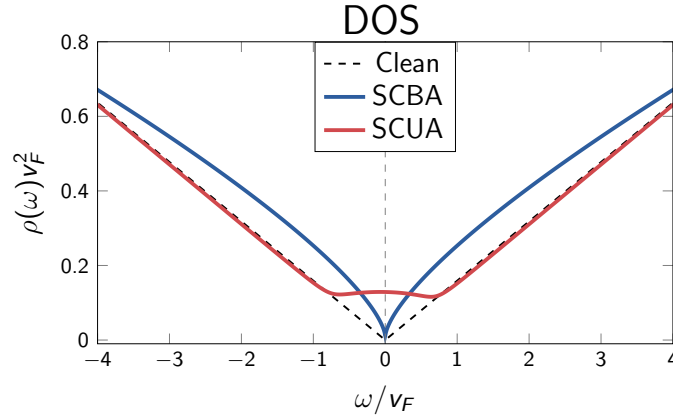


Fig. 2.2 Density of states of a topological insulator in the strong and weak scattering limit: The density of states for point-like scatterers in the self-consistent Born and unitary limits. The SCUA has an impurity dominated regime close to the Dirac point, indicated by a substantial increase in the self-energy – see chapter 5 for further details. (Same parameters as Fig. 2.1)

and density of states of the topological insulator surface. We saw that the impurities alter the density of states and this is highly dependent on the scattering strength of the impurity potentials. In the following chapters the concepts developed in this chapter will be at the center of many of our calculations.

Quasi-particle interference from the surface of topological insulators

In the previous chapter we saw how the surface states of topological insulators are protected from back-scattering. As our first foray into the world of disordered surface states, we consider a direct probe of this back-scattering protection: Quasi-particle interference (QPI) patterns measured by scanning tunneling microscopy (STM). STM measures the build up of standing-waves in the local density of states (LDOS) $\rho(\mathbf{r}, \omega)$ that occur in the vicinity of a defect. Due to the lack of back-scattering on the surface of a TI, these oscillations decay much faster than the same oscillations that occur near a defect in a standard 2DEG. We will show that the Fourier transform of the QPI signal is especially sensitive to even small amounts of back-scattering, making it an excellent direct experimental probe of a surface state's protection. At the end of this chapter we will see that a comparison of experiment and theory reveals a robust protection from back-scattering as well as allowing the extraction of the dispersion, scattering contributions, and the screening length of charged impurities. We therefore show that QPI is not only a good measure of back-scattering protection, but also a powerful tool for the characterisation of TI surface states..

3.1 LDOS due to an impurity

In the last chapter we saw that an impurity alters the energy landscape in its vicinity, breaking translational invariance and causing scattering between different momenta \mathbf{k} . In fact the new eigenstates attempt to minimise the additional energy resulting from the impurity and ensure the wave-function is reduced at the impurity position \mathbf{r}_i . For instance, consider a clean d -dimensional electron gas which has eigenstates that are plane-waves defined by a wave number \mathbf{k} and with eigenenergy $E_0(\mathbf{k}) = E_0(|\mathbf{k}|)$. If we now add a single point-like impurity potential located at \mathbf{r}_i , in order to reduce their energy at \mathbf{r}_i some of the new eigenstates of the clean system plus the impurity must form a superposition consisting of multiple $|\mathbf{k}|$, described the scattering perturbation theory discussed in the last chapter. In real space these new states are standing waves localised close to the impurity and decaying with distance $|\mathbf{r}|$ away from the impurity. It is clear that, as the dimension d increases, a more “efficient” superposition of \mathbf{k} states can be formed which is more highly localised about the impurity. Therefore suggesting that the effect of the impurity decays faster in real space with higher dimension. The best probe of this decay is the LDOS,

Chapter 3. Quasi-particle interference from the surface of topological insulators

which measures the square-modulus of the wave-function at an energy ω . It is defined

$$\rho(\mathbf{r}, \omega) = \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\mathbf{r}, E)|^2 \delta(\omega - \varepsilon_{\mathbf{k}}) \quad (3.1)$$

where $\psi_{\mathbf{k}}(\mathbf{r}, E)$ are the energy eigenstates of the full system, surface and impurity, in real space which have energies $\varepsilon_{\mathbf{k}}$. The LDOS of can be directly measured by scanning tunnelling microscopy (this will be discussed at the end of the chapter).

In a d -dimensional electron gas the oscillations that form around an impurity (in real-space at large distances $|\mathbf{r}|$) have the form [62]

$$\delta\rho(\mathbf{r}, \omega) = \rho(\mathbf{r}, \omega) - \rho_0(\omega) \sim \frac{\cos\left(\frac{2\omega}{v_F}|\mathbf{r} - \mathbf{r}_i| + \phi\right)}{|\mathbf{r}|^{d-1}}, \quad (3.2)$$

where ϕ is a constant phase and $\rho_0(\omega)$ the LDOS of the clean system. As our intuition in the above discussion suggested: Close to impurities standing waves form that are centered at the impurity and decay faster with increasing dimension d .

Since an electron gas is only occupied up to some Fermi-energy E_F , these oscillations in the LDOS also cause oscillations in the charge density, $n(\mathbf{r})$, which is simply the sum over all states below the Fermi-energy, such that

$$\delta n(\mathbf{r}) = \int_0^{E_F} d\omega \delta\rho(\mathbf{r}, \omega) \sim \frac{\cos(2k_F|\mathbf{r} - \mathbf{r}_i| + \phi)}{|\mathbf{r}|^d}, \quad (3.3)$$

with k_F the Fermi-wave vector. These oscillations in charge density are known as **Friedel oscillations** [71, 72]. Although in this chapter we will only be interested in oscillations of the LDOS, the slow decay of Friedel oscillations in low dimensions will be extremely important for the discussion of magnetoresistance in Chapter 8, where the application of a magnetic field causes these oscillations to become effectively one dimensional.

3.1.1 LDOS oscillations on a TI surface

Obviously the lack of backscattering in a Dirac material means an impurity has considerably less of an effect on the energy eigenstates of a Dirac material. It is important to note that, since we are considering a *local* probe, namely the LDOS, the probe does not self-average over impurity configurations (as, for example, is the case for the conductivity). In fact the impurity average $\langle \delta\rho(\mathbf{r}, \omega) \rangle_{\text{imp}}$ would be a constant with no \mathbf{r} dependence.¹ For the moment we are therefore really considering a single impurity at a fixed position \mathbf{r}_i (which we take to be the origin). The Hamiltonian of TI surface and impurity potential takes the form:

$$H = \sum_{\mathbf{k}} v_F \psi_{\mathbf{k}}^\dagger (\sigma_y k_x - \sigma_x k_y + m\sigma_z) \psi_{\mathbf{k}} + \sum_{\mathbf{k}, \mathbf{k}'} \psi_{\mathbf{k}}^\dagger u_{\mathbf{k}-\mathbf{k}'} \psi_{\mathbf{k}'}, \quad (3.4)$$

¹This is certainly not to say that these changes don't have an impact on any impurity averaged quantities or probes. As we will see in chapter 8 averages of higher order correlators such as $\langle u_0(\mathbf{r})\rho(\mathbf{r}', \omega) \rangle$ do not vanish and can have a substantial impact on, for example, the conductivity.

3.1. LDOS due to an impurity

where the spinor $\psi^\dagger(\mathbf{k}) = (c_{\mathbf{k},\uparrow}^\dagger, c_{\mathbf{k},\downarrow}^\dagger)$ and $u_{\mathbf{k}-\mathbf{k}'}$ is the Fourier transform of the impurity potential $u_0(\mathbf{r})$. For the moment we do not allow backscattering and so set $m = 0$, later we will show that enabling backscattering has a significant impact on the Fourier transformed LDOS (FT-LDOS).

In terms of the full real space Green's function, the LDOS is simply the $\rho(\mathbf{r}, \omega) = -\text{Im Tr } G(\mathbf{r}, \mathbf{r}, \omega)$ [73]. As we saw in the last chapter, however, the real space Green's function is normally a rather unwieldy object. It is much easier to consider the FT-LDOS in terms of the momentum space Green's function, which gives

$$\rho(\mathbf{q}, \omega) = -\frac{1}{\pi} \text{Im} \int \frac{d\mathbf{k}}{(2\pi)^2} \text{Tr } G(\mathbf{k} - \mathbf{q}, \omega). \quad (3.5)$$

Since we are currently only considering a single impurity the full Green's function is simply the sum of the free Green's function and scattering from the impurity [74, 75]

$$G(\mathbf{k} - \mathbf{q}, \omega) = \delta_{\mathbf{q},0} G^0(\mathbf{k}, \omega) + G^0(\mathbf{k}, \omega) T_{\mathbf{k},\mathbf{k}-\mathbf{q}}(\omega) G^0(\mathbf{k} - \mathbf{q}, \omega), \quad (3.6)$$

where $T_{\mathbf{k},\mathbf{k}-\mathbf{q}}$ is the the T-matrix from the last chapter.

3.1.2 Point-like scatterer

In chapter 2 we established the full T -matrix of a diagonal point-like scatterer $u_0(\mathbf{r}) = U\delta(\mathbf{r})\mathbb{1}$ which can be solved exactly since there is no-momentum dependence. The result was

$$T(\omega) = \frac{U\mathbb{1}}{1 + i\frac{U}{2v_F^2}\omega \ln\left(\frac{\omega}{v_F\Lambda}\right)}, \quad (3.7)$$

where Λ is the ultraviolet cut-off.

Due to the lack of momentum dependence in the T -matrix the full Green's function, Eq. (3.6), and the definition of the FT-LDOS, Eq. (3.5), can be used to analytically calculate the local density of states difference due to the point-like impurity, $\rho_\delta(\omega)$ [76]. To start we expand out the Green's function, such that

$$\delta\rho_S(\mathbf{q}, \omega) = -\frac{1}{\pi} \text{Im} \left\{ \text{Tr} \left\{ \int \frac{d\mathbf{k}}{(2\pi)^2} G^0(\mathbf{k}, \omega) T(\omega) G^0(\mathbf{k} + \mathbf{q}, \omega) \right\} \right\} = -\frac{1}{\pi} \text{Im} \{T(\omega) \Gamma(\mathbf{q}, \omega)\},$$

where

$$\Gamma(\mathbf{q}, \omega) = \int \frac{d\mathbf{k}}{(2\pi)^2} \frac{\omega^2 + v_f^2 k^2 + v_F^2 \mathbf{k} \cdot \mathbf{q}}{(\omega^2 - v_F^2 k^2)(\omega^2 - v_F^2 (\mathbf{k} + \mathbf{q})^2)}. \quad (3.8)$$

Such integrals as $\Gamma(\mathbf{q}, \omega)$, with a product in the denominator, can be calculated by moving to imaginary frequencies and using the Feynman parametrisation [75], which involves the

Chapter 3. Quasi-particle interference from the surface of topological insulators

introduction of an integral over the new variable x , the result is

$$\begin{aligned}\Gamma(\mathbf{q}, i\omega) &= \int_0^1 dx \int \frac{d^2\mathbf{k}}{(2\pi)^2} \frac{-\omega^2 + v_F^2 \mathbf{k}^2 - v_F^2 \mathbf{k} \cdot \mathbf{q}}{(-\omega^2 - v_F^2 \mathbf{k}^2 - v_F^2 (1-x) \mathbf{q}^2 + v_F^2 (1-x) 2\mathbf{k} \cdot \mathbf{q})^2} \\ &= \int_0^1 dx \int \frac{d^2\mathbf{k}}{(2\pi)^2} \frac{-\omega^2 + v_F^2 (k^2 - x(1-x) \mathbf{q}^2)}{(\omega^2 + v_F^2 \mathbf{k}^2 + x(1-x) v_F^2 \mathbf{q}^2)^2},\end{aligned}\quad (3.9)$$

where in the last line we made the shift of $\mathbf{k} \rightarrow \mathbf{k} + (1-x)\mathbf{q}$ and use of the fact several components in the denominator are symmetric about $x = 1/2$ and therefore cancel.

Defining where $\Delta^2 = \omega^2 + x(1-x)v_F^2 \mathbf{q}^2$ leaves

$$\Gamma(\mathbf{q}, i\omega) = \int_0^1 dx \int \frac{d^2\mathbf{k}}{(2\pi)^2} \frac{-\omega^2 + v_F^2 (\mathbf{k}^2 - x(1-x) \mathbf{q}^2)}{(\Delta^2 + v_F^2 \mathbf{k}^2)^2}. \quad (3.10)$$

We have transformed the integral such that we can trivially perform the angular integral, leaving only simple integrals over x and $k = |\mathbf{k}|$. The momentum integral results in,

$$\Gamma(\mathbf{q}, i\omega) = \int_0^1 dx \int_0^\Lambda \frac{dk}{2\pi} \frac{k}{k^2} \frac{-\omega^2 + v_F^2 (k^2 - x(1-x) \mathbf{q}^2)}{(\Delta^2 + v_F^2 k^2)^2} = \frac{1}{4\pi v_F^2} \int_0^1 dx \left(-2 - \ln \left(\frac{\Delta^2}{\Lambda^2} \right) \right), \quad (3.11)$$

where in the last line we used that $q \ll \Lambda$. Finally, performing the x integral and returning to real frequencies $\omega \rightarrow \omega + i\delta$ we have the analytic formula for the FT-LDOS difference,

$$\begin{aligned}\delta\rho_S(\mathbf{q}, \omega) &= -\frac{1}{\pi} \text{Im}\{T(\omega)\Gamma(\mathbf{q}, \omega)\} \\ &= \frac{1}{2\pi^2 v_F^2} \text{Im} \left\{ T(\omega) \left(-\ln \left(\frac{-\omega^2}{\Lambda^2} \right) - 2 \text{sgn}(\omega) g(\omega) \text{arcTan} \frac{1}{g(\omega)} \right) \right\},\end{aligned}\quad (3.12)$$

where $g(\omega) = \sqrt{-1 + \left(\frac{2\omega}{v_F q} \right)^2}$. Line cuts of the FT-LDOS – with and without broadening, see below – are plotted for strong and weak scatterers in Fig. 3.1 (red and blue lines, respectively). We see that $\delta\rho_\delta(\mathbf{q}, \omega)$ is continuous for all \mathbf{q} . It does, however, have a cusp at $|\mathbf{q}| = \frac{2\omega}{v_F}$, which is the backscattering vector connecting $\pm\mathbf{k}$ on the constant energy contour with energy ω . This cusp, rather than a divergence, at the nesting vector $|\mathbf{q}| = \frac{2\omega}{v_F}$ is directly related to the lack of backscattering on the TI surface [76]. It has the consequence the oscillations in the real space LDOS decay as $1/r^2$, such that in real space

$$\delta\rho_S(\mathbf{r}, \omega) \sim \frac{\cos \left(\frac{2\omega}{v_F} |\mathbf{r} - \mathbf{r}_i| + \phi \right)}{|\mathbf{r}|^2}. \quad (3.13)$$

This decay is much faster than the $1/r$ decay of a 2DEG.²

²It should be noted that the energy ω is measured from the Dirac point. The LDOS at the Fermi-energy is simply given by $\delta\rho_S(\mathbf{r}, \mu)$.

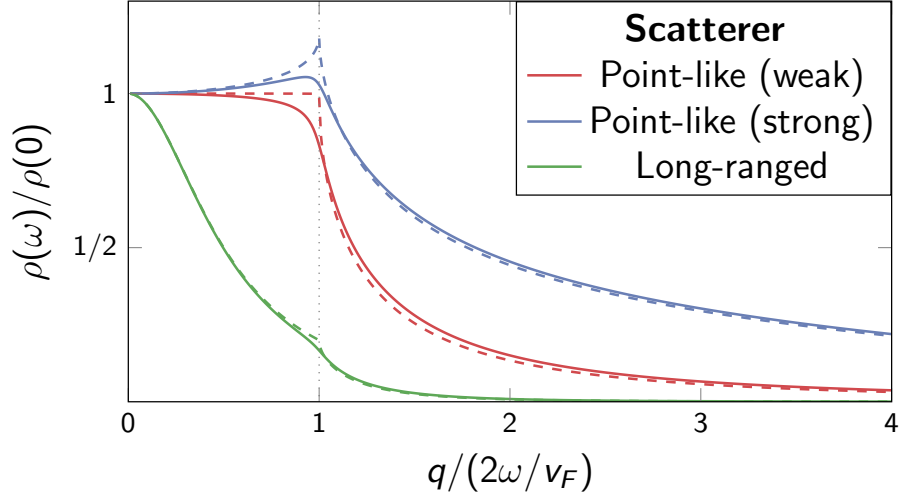


Fig. 3.1 **FT-LDOS line cuts for various scatterers**: Line-cuts of the FT-LDOS calculated for a weak point-like impurity (red), a strong unitary point-like impurity (blue), and long-ranged screened Coulomb impurity (green). Dashed lines show the FT-LDOS without any broadening effects. Solid lines show the broadened FT-LDOS with a finite self-energy. (Parameters used: $2\omega = 1$, $v_F = 1$, $\Sigma = i\omega/20$, $\kappa = \omega/v_F$, $\Lambda = 10\omega$).

3.1.3 Long-range scatterers

We now consider long-range scatterers, in particular the screened Coulomb potentials $U_{\mathbf{q}} = \frac{U}{q^2 + \kappa_0^2}$ where κ_0 is the screening wave-vector. The calculation of the full T -matrix for strong scatterers which are not point-like in general must be done numerically [75]. However, in the Born limit the T -matrix is simply $T_{\mathbf{k},\mathbf{q}}(\omega) = U_{\mathbf{q}}$, since the electrons only see the bare impurity potential. The FT-LDOS due to a weakly scattering Coulomb impurity is therefore,

$$\delta\rho_C(\mathbf{q}, \omega) = -\frac{1}{\pi} \text{Im}\{U_{\mathbf{q}}\Gamma(\mathbf{q}, \omega)\} \quad (3.14)$$

which is the same as the FT-LDOS due to a (weak) point-like impurity multiplied by the screened Coulomb potential (see Fig. 3.1).

3.1.4 Effect of quasi-particle broadening

In reality of course the free Green's function Eq. (2.7) is broadened by other scattering effects leading to decoherence and a finite self-energy $\Sigma(\omega)$, where the real-part can be absorbed in a redefinition of the chemical potential. We can include the self-energy in the FT-LDOS, $\rho(\omega)$, by the simple replacement $\omega \rightarrow \omega - \Sigma(\omega)$. Fig. 3.1 (solid lines) shows these broadened line-cuts of the FT-LDOS. The main impact of broadening is to smooth the divergence in the derivative of the FT-LDOS at $q = 2\omega/v_F$. An important consequence in real space is that the standing waves of the LDOS due to an impurity obtain an extra decay factor $\sim e^{-\frac{|\Sigma(\omega)|}{v_F}|r|}$. For weak scattering broadening has the impact that there is no sharp feature in the FT-LDOS at the nesting q -vector and, even for very strong-scatterers, any cusp is diminished (see Fig. 3.1).

3.1.5 Magnetic impurities

We now briefly consider a single point like spinful impurity of the form $U^\alpha(\mathbf{r}) = U\delta(\mathbf{r})\sigma^\alpha$, where for simplicity we only consider $\alpha = x$, so that the impurity's spin is always pointed in the x-direction. First weak scattering in the Born limit (as in [75]), such that $T^x(\omega) \approx U\sigma^x$, results in a FT-LDOS

$$\begin{aligned}\delta\rho_\sigma(\mathbf{q}, \omega) &= -\frac{U}{\pi} \text{Im} \left\{ \text{Tr} \left\{ \int \frac{d\mathbf{k}}{(2\pi)^2} G^0(\mathbf{k}, \omega) \sigma^x G^0(\mathbf{k} - \mathbf{q}, \omega) \right\} \right\} \\ &= -\frac{2U}{\pi} \text{Im} \left\{ \int \frac{d\mathbf{k}}{(2\pi)^2} \frac{2(q_y - 2k_y)}{(\omega^2 - v_F^2 k^2)(\omega^2 - v_F^2 (\mathbf{k} - \mathbf{q})^2)} \right\} \\ &= -\frac{2U}{\pi} \text{Im} \left\{ \int \frac{d\mathbf{k}}{(2\pi)^2} \frac{-4k_y}{(\omega^2 - v_F^2 (\mathbf{k} + \mathbf{q}/2)^2)(\omega^2 - v_F^2 (\mathbf{k} - \mathbf{q}/2)^2)} \right\} = 0,\end{aligned}\tag{3.15}$$

where in the last line we made the transformation $\mathbf{k} \rightarrow \mathbf{k} + \mathbf{q}/2$. This integral vanishes because it is odd under $\mathbf{k} \rightarrow -\mathbf{k}$. The same calculation holds regardless of spin direction α . Hence to first order in the Born expansion a magnetic impurity has no impact on the LDOS difference. It is important to note that in reality the spinful impurity causes two standing waves of electrons with opposite spin that are completely out of phase, the net result being no change in the LDOS, which is the sum over both spin components of the electrons. In contrast there would be an impact if we considered the *spin resolved* FT-LDOS which consider the contribution of a set spin-direction, such a quantity can be measured in STM by using a magnetic tip.

The calculation above also hold when considering higher order terms in the T -matrix expansion. For instance is all odd terms in $U^\alpha(\mathbf{r})$ will also vanish since they are proportional to σ^α and independent of momentum. On the other hand, even terms are proportional to the identity, this means such terms act in the same manner as point-like non-spinful scatterers and cannot cause any new features in the FT-LDOS at a given ω than we already found.

3.1.6 QPI due to massive Dirac fermions

Although only physically allowed when time reversal symmetry is broken, we now “turn on” the mass term $m\sigma_z$ in the Hamiltonian of the surface states, Eq. (3.4), this is to illustrate the effects of opening up a backscattering channel between surface states.

The calculation of the FT-LDOS proceeds identically as in the massless case and can be found in Appendix A. It turns out that the FT-LDOS is simply a sum of massless and massive terms

$$\rho(\omega) = \rho_0(\tilde{\omega}) + \rho_m(\tilde{\omega}),\tag{3.16}$$

3.2. QPI from $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ surfaces

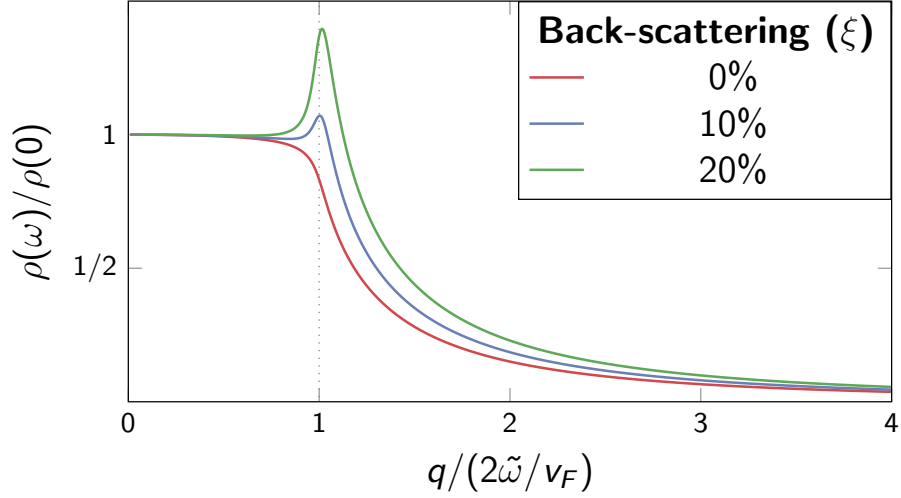


Fig. 3.2 **FT-LDOS line cuts with back-scattering**: Line-cuts of the (broadened) FT-LDOS calculated for a weak point-like impurity and finite back-scattering matrix element $\xi = |\langle \mathbf{k} | U | -\mathbf{k} \rangle|^2$, shown in percent. We see that even a small amount of back-scattering results in a noticeable peak at the nesting vector $q = 2\omega/v_F$. (Same parameters, with a finite self-energy as Fig. 3.1)

where $\tilde{\omega}^2 = \omega^2 - m^2$ is the energy measured from the bottom of the band – since m results in a gap of equal size – and the new term in the LDOS due to the mass is,

$$\rho_m(\tilde{\omega}) = -\frac{1}{\pi} \text{Im} \left\{ \frac{2 \left(\frac{2m}{v_F q} \right)^2 \text{arcTan} \frac{1}{g(\tilde{\omega})}}{g(\tilde{\omega})} \right\}, \quad (3.17)$$

where once again $g(\omega) = \sqrt{-1 + \left(\frac{2\omega}{v_F q} \right)^2}$. We see that the new term due to backscattering $\rho_m(\omega)$ has a $1/\omega$ divergence at $q = \frac{2\tilde{\omega}}{v_F}$ whereas our original $\rho_0(\omega)$ only had a cusp at the equivalent position. In practice what this term means that massive Dirac fermions have oscillations in the LDOS near an impurity that decay with a leading order $1/r^{d-1}$, just like their Schrödinger cousins.

Since a constant mass term results in a back-scattering matrix element $\xi = |\langle \mathbf{k} | U | -\mathbf{k} \rangle|^2$ that changes with energy it is better for our illustrative purpose in Fig. 3.2 to consider the reverse and set the back-scattering rate ξ constant for all ω , equivalent to setting the mass $m = \frac{\omega \sqrt{\xi}}{\sqrt{1-\xi}}$ (see previous chapter). It can be seen in Fig. 3.2 that, with just a small backscattering rate and even after broadening, the FT-LDOS has a significant peak at the nesting vector $q = 2\omega/v_F$ compared to Dirac fermions that are completely protected from back-scattering.

3.2 QPI from $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ surfaces

Having calculated the FT-LDOS theoretically we can now interpret the LDOS measured in real experiments. In particular we will discuss scanning tunneling spectroscopy (STS)

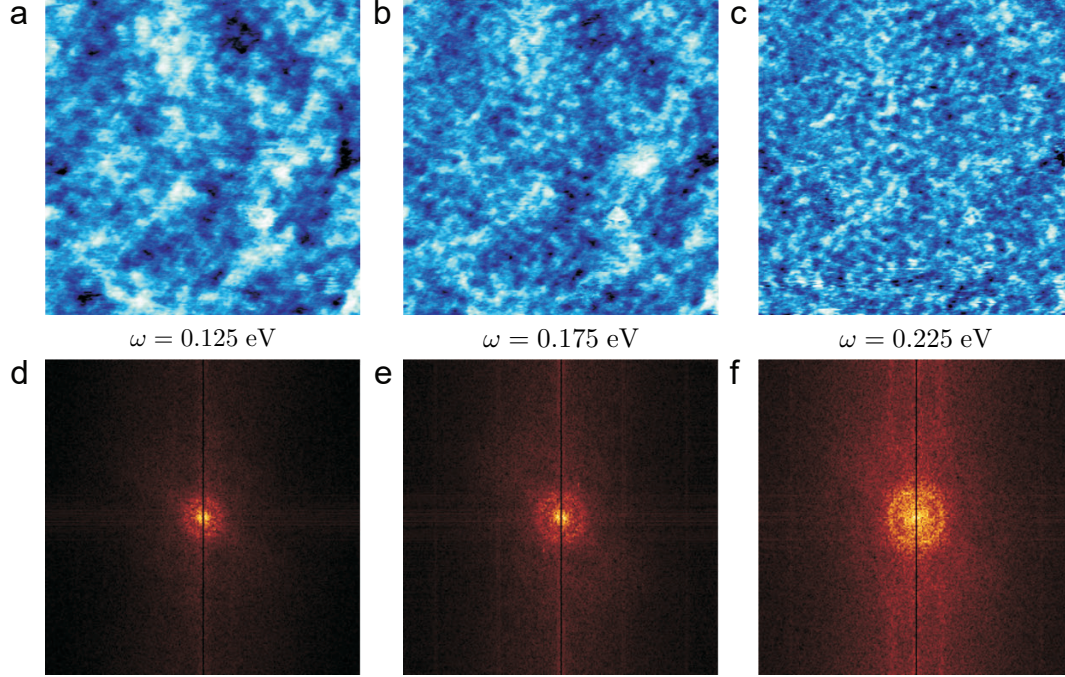


Fig. 3.3 **QPI from the surface of BSTS in real and momentum space:** **a - c)** Real space images of the LDOS at the $\omega=0.125$, 0.175 , and 0.225 eV, respectively. All images have an area $100 \text{ nm} \times 100 \text{ nm}$. Two types of fluctuations can be distinguished: Longer fluctuations $\sim 40 \text{ nm}$ and energy independent can be attributed to surface charged puddles (see chapter 1). Shorter oscillations reduce in wave-length with higher energy and are due to quasi-particle interference. **d - f)** Fourier transforms of the real-space data. A clear central peak is visible along with a ring which expands with increasing energy.

experiments performed on the bulk-insulating TI $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ (BSTS). Scanning tunnelling microscopy works by applying a bias voltage between the tip and the TI surface. This results in a tunnelling current between the surface and the tip which is proportional to the LDOS of the TI surface under the tip [73]. In this context the FT-LDOS is known as quasi-particle interference (QPI).

Impurities near or on the surface of BSTS not only act as scattering centers for quasiparticle interference (QPI) but also as charged impurities that can lead to the formation of charge puddles on the surface [77, 78]. Both surface charge puddles and QPI can be observed simultaneously by measuring scanning tunnelling spectroscopy (STS) map. In Fig. 3.3 a-c, one can clearly recognize two different fluctuation features: 1) The broader fluctuations of the order of 40 nm can be attributed to surface charge puddles (see [77] for further details on these puddles). 2) Regular, shorter wavelength oscillation are the standing waves in LDOS discussed above. This assignment is further verified by the fact that only the short wavelength show an energy dependence when increasing the bias voltage. The Fourier transforms of the real-space LDOS are shown in Fig. 3.3 d - f; in addition to a strong central peak, an isotropic ring that expands with energy is visible.

3.2. QPI from $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ surfaces

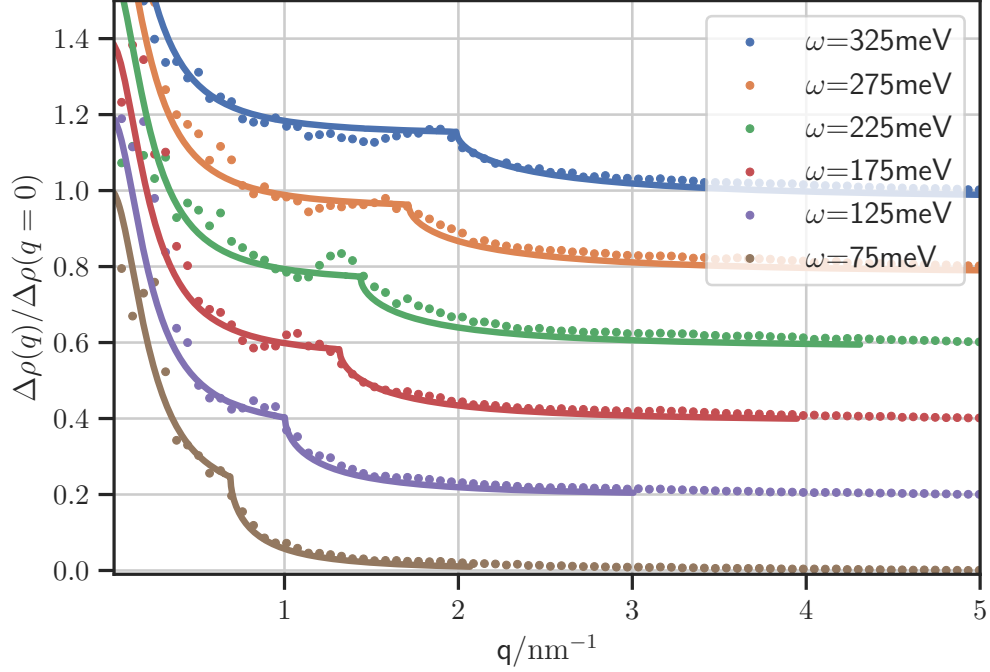


Fig. 3.4 **QPI line-scans, comparison of theory and experiment:** (Each energy is shifted by a constant, for clarity) Points: Angular averaged line-scans of the QPI measured on the surface of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$. Solid lines: FT-LDOS calculated theoretically in the weak scattering limit with no broadening from decoherence. Using ARPES data for the Fermi-velocity [79] the locations of kinks at $2\omega/v_F$ are therefore independent of theory fitting parameters. The height of the kink is set by an admixture of contributions to the FT-LDOS from long and short range impurities, we find $\frac{n_S}{n_L} = 0.17$ gives the best fits. The decay of the peak close to $q=0$ is set by the screening length, which we find to be $l \sim 9$ nm.

To compare theory and experiment we use the dispersion of BSTS as measured by ARPES [79], this sets entirely the location of the kink occurring at $2\omega/v_F$. We assume that the total LDOS change is an admixture of contributions from long-range and short range scatterers $\delta\rho(\mathbf{q}) = n_S\rho_S(\mathbf{q}) + n_L\rho_L(\mathbf{q})$, this sets the height of the kink in Fig. 3.4. The final fitting parameter is the screening length $l = 1/\kappa_0$, this sets the width of the long-range potential, which determines the decay of the FT-LDOS close to $q \approx 0$. Several of these line-scans are shown in Fig. 3.4. The same parameter set is used for all plots and no broadening is included. To be precise, we find that $\frac{n_S}{n_L} \approx 0.17$ and $l \sim 9$ nm produce the best fits. Fig. 3.4 shows the excellent agreement of experiment and theory. In particular the lack of a substantial peak in the QPI confirms the excellent protection from back-scattering on $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ surfaces.

Summary and outlook

In this chapter we used the T-matrix formalism developed in the previous chapter to calculate the oscillations in the LDOS induced by an impurity on the surface of a topological

Chapter 3. Quasi-particle interference from the surface of topological insulators

insulator. This relatively simple theoretical calculation was able to reproduce well quasi-particle interference measurements on the surface of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$, revealing the robust protection from back-scattering of a topological insulator's surface states. In the next chapter we will see the impact this protection has on its electrical conductivity before, in chapter 5, describing an experiment which allows us to break this protection.

Transport in Dirac matter

In this chapter we introduce for the first time one of the main themes of this thesis: Electronic transport. We saw in chapter 1 the ability of angle resolved photoemission (ARPES) to use the photoelectric effect to measure surface-bands $E(\mathbf{k})$ and diagnose a sample as a Dirac material. In chapter 3 we saw how another surface measurement, scanning tunnelling spectroscopy (STS), could directly experimentally probe a topological insulator's response to impurities through the LDOS. Unlike these probes, the conductivity of a material is only an indirect measure of its Dirac physics. Nonetheless, a material's conductivity is one of its most fundamental and useful properties. This is evinced by the fact physicists chose to name topological *insulators* in reference to their transport properties. As we will see in the coming chapter, from a theoretical perspective, transport in Dirac materials is especially interesting as a result of the topological protection of the Dirac states.

We will first discuss how conductivities can be calculated semi-classically using the Boltzmann equation. We then switch to the quantum equivalent, the Kubo formula, which will become our tool of choice in later chapters. As in chapter 2 we will use the specific example of a topological insulator's surface as a working example. In the second half of this chapter we will address some more experimental questions of importance for measuring surface transport. The first is the concept of compensation which enables the production of genuinely insulating topological insulators. Second we will discuss how it is possible to more finely vary the position the Fermi-energy by electrostatic gating.¹

Parts of this chapter are discussed in publication 1: “Planar Hall effect from the surface of topological insulators” Nat. Commun. 8, 1340 (2017). [70]

4.1 Semi-classical transport

When considering electron currents flowing in a metal the key material characteristic is the conductivity σ , this governs the size of the current density, $\mathbf{j} = \sigma \mathbf{E}$, formed when an electric field \mathbf{E} is applied. The Drude formula for conductivity considers electrons as classical ball bearings in a pinball machine [80]. It predicts the conductivity to be $\sigma = \frac{ne^2\tau_{\text{tr}}}{m^*}$, where n is the electron density, m^* its effective mass, and τ_{tr} is **transport scattering time**. The transport scattering time is a *different* quantity than the scattering time τ from chapter 2 since, as we will see, not all scattering events are relevant for transport. The Drude

¹We actually already saw in the first chapter that gating was required to push HgTe quantum wells into their bulk insulating phase and measure the e^2/h conductance from their edge states.[17]

formula is a nice phenomenological model for the conductivity, but the real theoretical challenge is to calculate τ_{tr} and correctly take into account the relevant *conservation laws*. We will see later that quantum mechanically this is a matter of *vertex corrections*,² but we will first find that we can ascertain many of the salient features of conductivities by using a semi-classical Boltzmann description. (The outline of this section follows the similar derivation for Schrödinger fermions in Ref. [66]).

4.1.1 The Boltzmann equation

The Boltzmann equation describes the evolution of a semi-classical wave packet with its dynamics encapsulated by the distribution function $f(\mathbf{r}, \mathbf{k}, t)$, where in our working example of a disordered TI surface the phase space coordinates are $\mathbf{r} = (r_x, r_y)$ and $\mathbf{k} = (k_x, k_y)$. In what follows we will assume the clean system is homogeneous and add a very dilute density of impurities. The dynamics of the distribution function are governed by the Boltzmann equation

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \frac{\partial \mathbf{k}}{\partial t} \cdot \nabla_{\mathbf{k}} f = I[f]. \quad (4.1)$$

Where the left-hand side is determined by the forces exerted on the electrons and the right-hand side, $I[f] = \left(\frac{df}{dt}\right)_{\text{coll}}$, is known as the *collision integral*, this takes into account the rate of change due to collisions of electrons with other objects in the system, for our purpose this means impurities. We also assume a constant homogeneous electric field \mathbf{E} such that there is no dependence of the distribution function on position \mathbf{r} or on time. This means the first two terms on the left-hand-side can be dropped and $f(\mathbf{r}, \mathbf{k}, t) \equiv f(\mathbf{k})$. If in addition no magnetic field perpendicular to the TI surface is then present the remaining term, the acceleration, is equal to the forces on the electron such that $\frac{\partial \mathbf{k}}{\partial t} = -e\mathbf{E}$, since the only external force is the electric field.

To determine the right-hand side of the Boltzmann equation we will assume that the electric field is weak, such that the leading change in the distribution function is its linear response around an equilibrium position $f(\mathbf{k}) \approx f_{\text{eq}}(\mathbf{k}) + f^{(1)}(\mathbf{k})$. The equilibrium is of course just the Fermi-Dirac distribution of the electrons $f_0(\mathbf{k}) = n_F(\varepsilon(\mathbf{k}) - \mu)$, with μ the chemical potential. In general collisions have the effect of relaxing the system back to its equilibrium distribution $f_0(\mathbf{k})$ and we can make the **relaxation time approximation** [80]

$$I[f] = -\frac{f(\mathbf{k}) - f_{\text{eq}}(\mathbf{k})}{\tau_{\text{tr}}(\mathbf{k})} = -\frac{f^{(1)}(\mathbf{k})}{\tau_{\text{tr}}(\mathbf{k})} \quad (4.2)$$

where $\tau_{\text{tr}}(\mathbf{k})$ is the scattering rate relevant for transport, as appears in the Drude conductivity. Within this approximation and inserting the electric field the Boltzmann equation of Eq. (4.1) is a relation between $f^{(1)}$ and f , such that

$$e\mathbf{E} \cdot \nabla_{\mathbf{k}} f(\mathbf{k}) \approx e\mathbf{E} \cdot \nabla_{\mathbf{k}} f_{\text{eq}}(\mathbf{k}) = \frac{f^{(1)}(\mathbf{k})}{\tau_{\text{tr}}(\mathbf{k})}, \quad (4.3)$$

²Mahan sums this up nicely at the start of his chapter on DC conductivities [66]: “The present chapter is really a chapter about vertex corrections.”

4.1. Semi-classical transport

where on the left hand side we have used that $f^{(1)}$ is a small deviation from the equilibrium position. Using Eq. 4.3 as an equation for $f^{(1)}$, we obtain the relaxation time approximation to full distribution function

$$f(\mathbf{k}) \approx f_0(\mathbf{k}) + e\tau_{\text{tr}}(\mathbf{k})\mathbf{E} \cdot \nabla_{\mathbf{k}}f_{\text{eq}}(\mathbf{k}) = f_0(\mathbf{k}) + ev_F\tau_{\text{tr}}(\mathbf{k})\mathbf{E} \cdot \hat{\mathbf{k}} \frac{\partial f_0(\mathbf{k})}{\partial \varepsilon_{\mathbf{k}}}. \quad (4.4)$$

In the final equality, to turn the momentum \mathbf{k} derivative into an easier to handle energy derivative we introduced the velocity $\mathbf{v}_{\mathbf{k}}$, which is especially simple in the case of a topological insulator³ since $\mathbf{v}(\mathbf{k}) = \nabla_{\mathbf{k}}\varepsilon(\mathbf{k}) = v_F\hat{\mathbf{k}}$.

We see that the impact of the electric field in Eq. 4.4 on the equilibrium distribution function is to shift the distribution parallel to the electric field resulting in slightly more electrons on the one side of the Fermi-surface compared to the other. The net result is that a current flows with density \mathbf{j} . In particular we can calculate this current in terms of the averaged velocity $\langle \mathbf{v} \rangle$, which is averaged over the distribution function f such that

$$\mathbf{j} = -e\langle \mathbf{v} \rangle = -e \int \frac{d^2\mathbf{k}}{(2\pi)^2} \mathbf{v}(\mathbf{k})(f(\mathbf{k}) - f_0(\mathbf{k})) = -e^2v_F^2 \int \frac{d^2\mathbf{k}}{(2\pi)^2} \tau_{\text{tr}}(\mathbf{k})\hat{\mathbf{k}}(\mathbf{E} \cdot \hat{\mathbf{k}}) \frac{\partial f_0(\mathbf{k})}{\partial \varepsilon_{\mathbf{k}}}. \quad (4.5)$$

The isotropic Fermi-surface close to the Dirac point and – in this chapter – isotropic scatterers mean that all quantities are functions of $k = |\mathbf{k}|$ only. This enables us to replace the integral over $\frac{d^2\mathbf{k}}{(2\pi)^2}$ with an integral over $\rho(\varepsilon_{\mathbf{k}})d\varepsilon_{\mathbf{k}}$, where $\rho(\varepsilon_{\mathbf{k}})$ is the density of states. We can also take the low temperature limit such that the derivative of the Fermi-Dirac distribution is $\frac{\partial f_0(\mathbf{k})}{\partial \varepsilon_{\mathbf{k}}} \approx -\delta(\varepsilon_{\mathbf{k}} - \mu)$. Including these aspects in Eq. (4.5) we obtain $j_{\parallel} = \sigma(\mu)E_{\parallel}$, where j_{\parallel} is the component of the current density parallel to the field and the conductivity is

$$\sigma(\mu) = \frac{1}{2}e^2v_F^2\rho(\mu)\tau_{\text{tr}}(k_F), \quad (4.6)$$

with $k_F = \mu/v_F$ the Fermi-momentum. This is a generalised form of the Drude formula, for which we still need to calculate the transport scattering time $\tau_{\text{tr}}(k_F)$ in order to determine the conductivity of our topological insulator's surface.

4.1.2 The transport scattering time

Fortunately our discussion of the T-matrix in chapter 2 and the above derivation of the distribution function $f(\mathbf{k})$ has provided all the tools necessary to calculate the transport scattering time $\tau_{\text{tr}}(k_F)$. By definition the collision integral $I[f(\mathbf{k})]$ is the rate of scattering into a state \mathbf{k} minus the rate of out-scattering. We know from chapter 2 that such a rate is governed by the T-matrix's matrix elements $T_{\mathbf{k}\mathbf{k}'} \equiv \langle \mathbf{k} | \mathbf{T} | \mathbf{k}' \rangle$, where \mathbf{k}' are the states being scattered to or from \mathbf{k} . The collision integral in the relaxation time approximation is

³Assuming we are close enough to the Dirac point so that the Fermi-surface is isotropic and we can neglect any additional hexagonal warping terms.

therefore

$$\begin{aligned} -I[f] &= 2\pi n_{\text{imp}} \int \frac{d^2 \mathbf{k}'}{(2\pi)^2} \delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}'}) \left[|T_{\mathbf{k}\mathbf{k}'}|^2 f(\mathbf{k})(1 - f(\mathbf{k}')) - |T_{\mathbf{k}'\mathbf{k}}|^2 f(\mathbf{k}')(1 - f(\mathbf{k})) \right] \\ &= \frac{f^{(1)}(\mathbf{k})}{\tau_{\text{tr}}(\mathbf{k})} = 2\pi n_{\text{imp}} \int \frac{d^2 \mathbf{k}'}{(2\pi)^2} \delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}'}) |T_{\mathbf{k}'\mathbf{k}}|^2 [f(\mathbf{k}) - f(\mathbf{k}')], \end{aligned} \quad (4.7)$$

where energy conservation is enforced by $\delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}'})$ because impurities are elastic scatterers. Inserting our expansion of the distribution function $f(\mathbf{k})$ from Eq. (4.4) we see that all factors from the distribution functions on both sides of Eq. (4.7) cancel besides $\hat{\mathbf{E}} \cdot \hat{\mathbf{k}} = \cos(\phi)$ and $\hat{\mathbf{E}} \cdot \hat{\mathbf{k}}' = \cos(\theta - \phi)$, where $\theta = \hat{\mathbf{k}} \cdot \hat{\mathbf{k}}' = \theta_{\mathbf{k}} - \theta_{\mathbf{k}'}$ is the angle between the two momenta. Due to the Fermi-surface and scatterer's isotropy τ_{tr} is only a function of $k = |\mathbf{k}|$ and so without a loss of generality we can choose $\phi = 0$. The resulting equation for the transport scattering time is

$$\frac{1}{\tau_{\text{tr}}(k)} = 2\pi n_{\text{imp}} \int \frac{d^2 \mathbf{k}'}{(2\pi)^2} |T_{\mathbf{k}'\mathbf{k}}|^2 \delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}'}) (1 - \cos(\theta)). \quad (4.8)$$

The factor $1 - \cos(\theta)$ distinguishes the transport scattering time from the scattering time. It takes into account that scattering events which do not strongly change the angle of the particle have much less of an impact than larger angle scattering events, such as back-scattering. The underlying reason for this term is that we need to correctly take into account conservation laws – as classically enforced by the current continuity equation $e \frac{\partial \mathbf{j}}{\partial t} + \nabla \cdot \mathbf{j} = 0$ – which would be violated in its absence. In particular, scattering into the “beam” of electrons – i.e. parallel to \mathbf{E} – does not lead to a decay of current but, if this term were absent, such scattering events would have falsely diminished the conductivity.

This factor is particularly important for the Dirac surface states of a topological insulator where, as we saw in chapter 2, at all orders in the T-matrix back-scattering is forbidden and large angle scattering suppressed. If for simplicity we consider point-like scatterers in the Born-limit, where we saw that the T-matrix gives $|\langle \mathbf{k} | \mathbf{T} | \mathbf{k}' \rangle|^2 = u_0^2 \cos^2\left(\frac{\theta}{2}\right)$ we find that the transport scattering time is [81]

$$\frac{1}{\tau_{\text{tr}}(k_F)} = u_0^2 n_{\text{imp}} \rho(\mu) \int d\theta \cos^2\left(\frac{\theta}{2}\right) (1 - \cos(\theta)) = \frac{\pi u_0^2 n_{\text{imp}}}{2} \rho(\mu). \quad (4.9)$$

This is *four times* smaller than if there was no topological protection factor from $|T_{\mathbf{k}'\mathbf{k}}|^2$, evincing the fact that protection from back-scattering has a considerable increase on the mobility of Dirac states. The transport scattering time is also two times smaller than the bare scattering time τ , which does not include the factor $1 - \cos(\theta)$. Inserting this into the conductivity Eq. (4.6) we find that the semi-classical conductivity for point-like scatterers in the Born limit is a constant [65] – since the density of states factors cancel – such that

$$\sigma^{\text{Boltz}} = \frac{e^2 v_F^2 \hbar}{\pi u_0^2 n_{\text{imp}}}, \quad (4.10)$$

where we have reinserted \hbar which was set to 1 in the rest of the derivation. This result is only valid for chemical potentials far from the Dirac point, we will see in the next section how a more accurate description in the vicinity of the Dirac point can be achieved. The conductivity is only constant when higher order terms in the T-matrix can be neglected, in particular for strong scatterers the conductivity develops a dependence $\sim \mu\rho(\mu)$, we will discuss a scenario with strong scattering in the next chapter.

4.2 Linear response

In the previous section we saw several important features of the conductivity of a topological insulator:⁴ 1) There is a difference between transport scattering time τ_{tr} and scattering time τ . This was because small angle scattering events have little impact on the conductivity. 2) Topological protection provides also increased mobility due to the lack of back-scattering or large angle scattering.

We can put these facts on a more concrete quantum mechanical footing by using linear response theory.⁵ In general linear response allows us to calculate the change in expectation of some operator $\langle \mathcal{O}(t) \rangle$ resulting from the application of a weak external field \mathbf{F} coupled to a force $f(t)$. For instance in our present discussion this will be the current operator in the α -direction, \hat{j}_α , and the gauge field \mathbf{A} . The change in the expectation of the operator \mathcal{O} due to the field \mathbf{F} is given by [62]

$$\delta\langle \mathcal{O}(t) \rangle = \frac{i}{\hbar} \int_{-\infty}^{\infty} dt' \theta(t-t') \langle [\mathcal{O}(t), \mathcal{O}(t')] \rangle f(t'), \quad (4.11)$$

where $\theta(x)$ is the Heaviside theta. The object $\frac{i}{\hbar} \theta(t-t') \langle [\mathcal{O}(t), \mathcal{O}(t')] \rangle$ is known as the *response function*, which is just a two-point correlation function. For us the important response function will be the longitudinal current-current correlation function

$$\pi_{\alpha\alpha}(t) = \frac{i}{\hbar} \theta(t-t') \text{Tr} \left\{ \langle [j_\alpha(t), j_\alpha(t')] \rangle \right\}, \quad (4.12)$$

where the trace runs over all quantum numbers (i.e. momentum and spin). There is a small subtlety here in now calculating the conductivity, this is because the electric field in terms of the gauge field \mathbf{A} is given by $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}$ or equivalently in momentum space $\mathbf{A}(\mathbf{k}) = -i\mathbf{E}(\mathbf{k})/\omega$. After also Fourier transforming the current-current correlator in energy and momentum $\pi_{\alpha\alpha}(\Omega)$, the correct (retarded) response function for the static (DC) longitudinal conductivity is [66]

$$\sigma_{\alpha\alpha} = - \lim_{\Omega \rightarrow 0} \left(\frac{\pi_{\alpha\alpha}(\Omega)}{i\Omega} \right) = - \lim_{\Omega \rightarrow 0} \left(\frac{\text{Im}[\pi_{\alpha\alpha}(\Omega)]}{\Omega} \right), \quad (4.13)$$

⁴Much of the discussion also extends to graphene where there are twice the number of Dirac points. That said, in graphene, one has to be more careful about inter-nodal scatterings, which only occur if impurities are short-ranged, these are not protected from back-scattering.

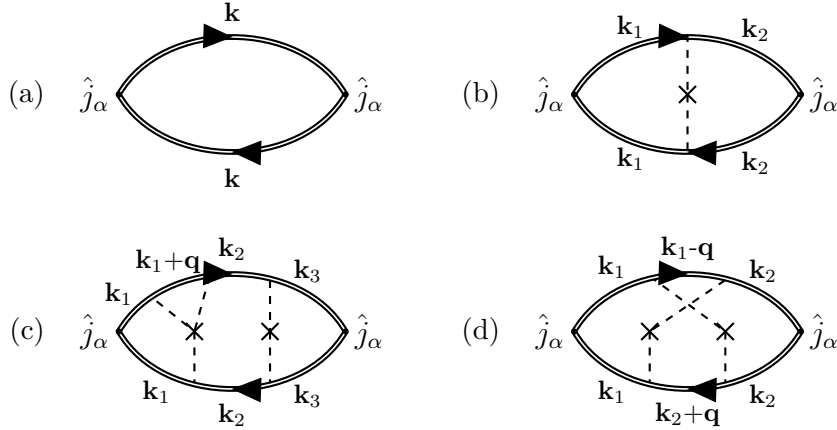
⁵A derivation of the Kubo formula and discussion of response functions can be found in e.g. Ref. [62].

where in the last inequality we made use of the fact the current is entirely real in the DC limit.

Diagrammatically⁶ the current-current correlation function can be represented as

$$\pi_{\alpha\alpha}(i\Omega) = \hat{j}_\alpha \begin{array}{c} \xrightarrow{i\omega_m, \mathbf{k}} \\ \xleftarrow{i\omega_m + i\Omega, \mathbf{k}} \end{array} \hat{j}_\alpha, \quad (4.14)$$

where the black triangle indicates the inclusion of vertex corrections and ω_m are the Matsubara frequencies, see below. Some diagrams contributing to this expansion are for instance



where the impurity lines have been averaged over and we have dropped the frequency index since scattering is elastic, such that the top and bottom Green's function lines always contain $i\omega_m$ and $i\omega_m + i\Omega$, respectively. The first diagram is the “bare bubble” contribution and the remaining diagrams are vertex corrections, these link both sides of the bubble with at least one scattering line. It may appear that the *vertex correction* terms are of a higher order and therefore of less relevance. This is actually not always the case, the reason can be seen in our equation for the semi-classical conductivity, Eq. (4.10), which we found to be *inversely* proportional to the impurity density n_{imp} . As such, *all* terms inversely proportional to n_{imp} are of equal relevance to the conductivity.

We can determine which terms contribute at order $1/n_{\text{imp}}$ by simple counting argument:

- Each scattering center with at least one scattering line connecting two sides of the bubble contributes a factor $\sim n_{\text{imp}}$.

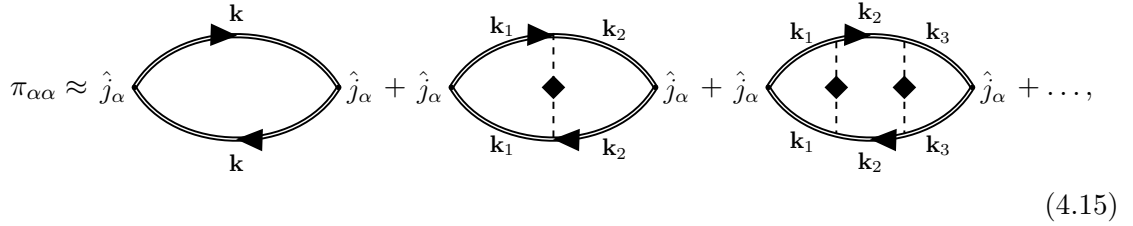
⁶We must use the full Green's function \mathcal{G} – or at the very least include *some* self-energy effects – since a system without a mechanism for current decay could not be resistive.

4.2. Linear response

- The leading terms in the self-energy $\Sigma(\omega)$ are proportional to n_{imp} so that, after integrating over the momentum, any pair of Green's function involving the same momentum, \mathbf{k}_i , will contribute a factor $\sim 1/n_{\text{imp}}$.

If we now use these criteria on these diagrams, we see that the bare bubble, (a), has only a pair of Green's functions and so is of order $1/n_{\text{imp}}$, as expected. Similarly diagrams (b) and (c) are also of order $1/n_{\text{imp}}$ because they only have integrals over pairs of Green's function and these cancel the additional n_{imp} factors from the increased number of scattering centers. In contrast diagram (d) is of a higher order since the crossing of the scattering lines reduces the number of completely free momentum integrals to two whilst the two scatterers still contribute a factor n_{imp} each, this diagram can therefore be neglected for small impurity densities $n_{\text{imp}} \rightarrow 0$.

Summing all the diagrams of order $1/n_{\text{imp}}$ we find the low density approximation to the current-current correlation function is



$$\pi_{\alpha\alpha} \approx \hat{j}_\alpha \text{ (bare bubble) } \hat{j}_\alpha + \hat{j}_\alpha \text{ (one T-matrix) } \hat{j}_\alpha + \hat{j}_\alpha \text{ (two T-matrices) } \hat{j}_\alpha + \dots, \quad (4.15)$$

where the diamond indicates the full T-matrix, as defined in chapter 2. These are known as **ladder diagrams** since each scatterer creates a “rung” between the two Green's functions. We will see that these ladder diagrams are extremely important for enforcing the conservation laws that led to the $1 - \cos(\theta)$ factor in the Boltzmann equation [66].

4.2.1 Bare bubble

To calculate the bare bubble contribution to the conductivity we need to trace over all quantum numbers: The Matsubara frequencies $\omega_m = \pi(2m + 1)/\beta$ with $m \in \mathbb{Z}$, and the momentum \mathbf{k} , such that

$$\pi_{\alpha\alpha}^{(0)}(i\omega) = \frac{1}{\beta} \sum_{\omega_m} \int \frac{d^2\mathbf{k}}{(2\pi)^2} [\hat{j}_\alpha \mathcal{G}(i\omega_m + i\omega, \mathbf{k}) \hat{j}_\alpha \mathcal{G}(i\omega_m, \mathbf{k})], \quad (4.16)$$

where the current operator $\hat{j}_\alpha = e \frac{\partial H}{\partial \mathbf{k}} = ev_F(\sigma_y, -\sigma_y)_\alpha$.

Summing over the Matsubara frequencies and performing analytic continuation such that $i\omega \rightarrow \omega + i\delta$ (see Appendix B for details) we obtain

$$\begin{aligned} \pi_{\alpha\alpha}^{(0)}(\Omega) = & e^2 v_F^2 \int \frac{d\omega}{2\pi i} \int \frac{d^2\mathbf{k}}{(2\pi)^2} \left(n_F(\omega) \text{Tr} \left\{ \sigma_\alpha \mathcal{G}(\mathbf{k}, \omega + \Omega + i\delta) \sigma_\alpha (\mathcal{G}(\mathbf{k}, \omega + i\delta) - \mathcal{G}(\mathbf{k}, \omega - i\delta)) \right\} \right. \\ & \left. + n_F(\omega - \Omega) \text{Tr} \left\{ \sigma_\alpha (\mathcal{G}(\mathbf{k}, \omega + i\delta) - \mathcal{G}(\mathbf{k}, \omega - i\delta)) \sigma_\alpha \mathcal{G}(\mathbf{k}, \omega - \Omega - i\delta) \right\} \right), \quad (4.17) \end{aligned}$$

where $\delta \rightarrow 0^+$ indicates the retarded Green's function $\mathcal{G}(\mathbf{k}, \omega)$ for $+i\delta$ and advanced Green's function $\mathcal{G}^\dagger(\mathbf{k}, \omega)$ for $-i\delta$.⁷ Taking the DC limit of this response function in Eq. (4.13) for the longitudinal conductivity we obtain

$$\sigma_{\alpha\alpha}^{(0)} = e^2 v_F^2 \text{Im} \int \frac{d\omega}{2\pi i} \int \frac{d^2\mathbf{k}}{(2\pi)^2} \frac{\partial n_F(\omega)}{\partial \omega} \text{Tr} \left\{ \sigma_\alpha \mathcal{G}(\mathbf{k}, \omega) \sigma_\alpha \mathcal{G}^\dagger(\mathbf{k}, \omega) \right\}, \quad (4.18)$$

$$= \frac{e^2 v_F^2}{2\pi} \text{Re} \int \frac{d^2\mathbf{k}}{(2\pi)^2} \text{Tr} \left\{ \sigma_\alpha \mathcal{G}(\mathbf{k}, \mu) \sigma_\alpha \mathcal{G}^\dagger(\mathbf{k}, \mu) \right\}, \quad (4.19)$$

where the second line is valid for low temperatures such that $\frac{\partial n_F(\omega)}{\partial \omega} \approx -\delta(\omega - \mu)$. We have dropped terms proportional to $\sigma_\alpha \mathcal{G} \sigma_\alpha \mathcal{G}$ and $\sigma_\alpha \mathcal{G}^\dagger \sigma_\alpha \mathcal{G}^\dagger$ which only contribute a small constant $\sim e^2/h$ conductivity that is negligible for all chemical potentials μ , apart from very close to the Dirac point, $\mu \approx 0$, where the non-crossing approximation anyway breaks down.

Performing the trace and integral over momentum space angle in Eq. (5.18), we are left with only an easy to perform integral over $k = |\mathbf{k}|$, such that

$$\sigma_{\alpha\alpha}^{(0)} = \frac{e^2 v_F^2}{2\pi} \int \frac{dk}{2\pi} \frac{2k(\mu^2 + (\text{Im}\Sigma)^2)}{|(\mu - i\text{Im}\Sigma)^2 - v_F^2 k^2|^2} = e^2 v_F^2 \frac{\mu^2 + (\text{Im}\Sigma)^2}{8\pi|\mu||\text{Im}\Sigma|}, \quad (4.20)$$

where we have absorbed the real part of the self-energy $\sim u_0 n_{\text{imp}}$ into the definition of μ . Since the density of impurities is very low and we are far from the Dirac point $\mu \gg \text{Im}\Sigma$ and we can neglect the second term in the numerator for our present discussion. Using the Born approximation – i.e. Eq. (2.25) from chapter 2 without the self-consistency – the self energy is $\Sigma_B(\mu) = \frac{-in_{\text{imp}} u_0^2 |\mu|}{4v_F^2}$ and we find the longitudinal conductivity is

$$\sigma_{\alpha\alpha}^{(0)} = \frac{e^2 v_F^2 \hbar}{2\pi u_0^2 n_{\text{imp}}} = \frac{1}{2} \sigma^{\text{Boltz}}. \quad (4.21)$$

We have succeeded in obtaining the same conductivity as we did using the Boltzmann equation, Eq. (4.10), but with an unwanted extra factor of 1/2. This corresponds exactly to the factor 2 contributed by the $1 - \cos\theta$ in the definition of the transport scattering time, that we argued was related to conservation laws. It is clear if we want to get the correct conductivity our calculation needs to be modified in a manner that includes these conservation laws.

4.2.2 Vertex corrections

The incorrect factor in comparison to the Boltzmann equation might be surprising. In particular, in every diagram that we calculated in chapter 2 we obeyed all conservation laws of particle number, momentum (on average), and energy at each vertex. The bare bubble itself also obeys these conservation laws at all vertices. What is missing, however,

⁷We will keep the conjugate transpose symbol \mathcal{G}^\dagger , rather than switch to a notation like \mathcal{G}^A , to remind us the Green's function is a matrix of Pauli matrices.

4.2. Linear response

are the vertex corrections from ladder diagrams which we argued also contribute the same order $1/n_{\text{imp}}$ to the conductivity as the bare bubble.

The ladder diagrams create a series of contributions to the conductivity $\sigma_{\alpha\alpha}^{(n)}$ with the n -th contribution having n T-matrix lines connecting the two Green's functions. In general these have the form

$$\sigma_{\alpha\alpha}^{(n)} = \frac{e^2 v_F^2}{2\pi} \text{Re} \int \left(\prod_{i=0}^n \frac{d^2 \mathbf{k}_i}{(2\pi)^2} \right) \text{Tr} \left\{ n_{\text{imp}}^n \sigma_\alpha \mathcal{G}(\mathbf{k}_0, \mu) \mathcal{V}_n \sigma_\alpha \mathcal{V}_n^\dagger \mathcal{G}^\dagger(\mathbf{k}_0, \mu) \right\}, \quad (4.22)$$

where the 2×2 matrix $\mathcal{V}_n = \prod_{i=1}^n (T(\mu) \mathcal{G}(\mathbf{k}_i, \mu))^n$ is due to the vertex corrections. Due to the fact the pair of Green's functions with \mathbf{k}_i appear at different points in the string of matrix products in Eq. (4.22), we cannot trivially perform each momentum integral separately. Fortunately the trace over the matrices enables us to redefine the same product in terms of 4×4 matrices of Kronecker products, such that (see Appendix B)

$$\mathcal{M} = \int \frac{d^2 \mathbf{k}_i}{(2\pi)^2} \mathcal{G}(\mathbf{k}_i, \mu) \otimes \mathcal{G}^\dagger(\mathbf{k}_i, \mu) \quad \text{and} \quad \mathcal{T} = T(\mu) \otimes T(\mu) \quad (4.23)$$

and by mapping the Pauli matrices to vectors, for instance $\sigma_x \rightarrow (1, 0, 0, 1)$. Within this notation the n -th contribution to the conductivity is

$$\sigma_{\alpha\alpha}^{(n)} = \frac{e^2 v_F^2}{2\pi} \text{Re} \left\{ n_{\text{imp}}^n \sigma_\alpha \mathcal{M} (\mathcal{T} \mathcal{M})^n \sigma_\alpha \right\}. \quad (4.24)$$

To obtain the full conductivity we need to sum over all ladder diagrams

$$\begin{aligned} \sigma_{\alpha\alpha} &= \sum_{n=0}^{\infty} \sigma_{\alpha\alpha}^{(n)} = \frac{e^2 v_F^2}{2\pi} \text{Re} \left\{ \sigma_\alpha \mathcal{M} \left(\mathbb{1} + n_{\text{imp}} \mathcal{T} \mathcal{M} + (n_{\text{imp}} \mathcal{T} \mathcal{M})^2 + \dots \right) \sigma_\alpha \right\} \\ &= \frac{e^2 v_F^2}{2\pi} \text{Re} \left\{ \sigma_\alpha \mathcal{M} (\mathbb{1} - n_{\text{imp}} \mathcal{T} \mathcal{M})^{-1} \sigma_\alpha \right\}, \end{aligned} \quad (4.25)$$

where in the second line we have used the geometric sum rule for matrices. The integral appearing on the diagonals of \mathcal{M} is the same as in Eq. (4.20) and all other terms in \mathcal{M} are either zero or do not contribute after taking the left and right vector product with σ_α . In particular in the Born limit the T-matrix is replaced by the bare potential such that $\mathcal{T} = u_0^2 \mathbb{1}$ and the diagonals of \mathcal{M} are given by $\mathcal{M}_{ii} = \frac{\mu^2}{8|\mu| |\text{Im} \Sigma_B|} = \frac{1}{2u_0^2 n_{\text{imp}}}$. The term governing the vertex corrections is therefore the diagonal $n_{\text{imp}} (\mathcal{T} \mathcal{M})_{ii} = \frac{1}{2}$, inserting this into Eq. (4.25) we obtain for the full conductivity, including vertex corrections,

$$\sigma_{\alpha\alpha} = \frac{e^2 v_F^2}{2\pi} \text{Re} \left\{ 2 \sigma_\alpha \mathcal{M} \sigma_\alpha \right\} = 2 \sigma_{\alpha\alpha}^{(0)} = \frac{e^2 v_F^2 \hbar}{\pi u_0^2 n_{\text{imp}}}. \quad (4.26)$$

Which is now exactly the same as we found using the Boltzmann equation in Eq. (4.10), including the factors due to the conservation laws.

We have managed to find the conductivity of a topological insulator's surface via two separate routes: Using the Boltzmann equation and using linear response theory. Since

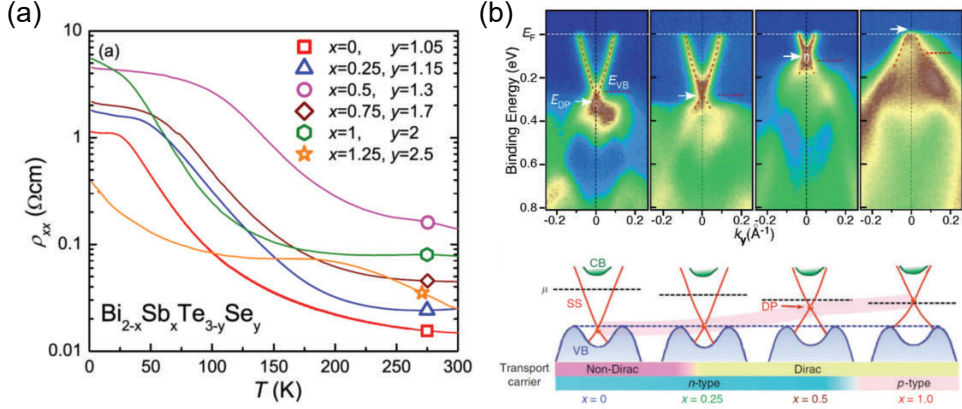


Fig. 4.1 **Compensation in 3d topological insulators:** (a) Temperature dependence of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ resistivity for various doping compositions. By introducing dopants one is able to achieve surface dominated transport. (b) Top: Band structure of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ as measured by ARPES for $(x, y) = (0, 1), (0.25, 1.15), (0.5, 1.3),$ and $(1, 2)$. Bottom: Schematic of the same band structure, showing the change in location of the Dirac point as a function of doping. (a: Reproduced from "Optimizing solid solutions to approach the intrinsic topological insulator regime." PRB 84, 165311 (2011). Ren Z., et al. [54] b: Reproduced from "Tunable Dirac cone in the topological insulator $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$." Nat.Comm. 3, 636 (2012). Arakane, T et al. [82]).

scattering in the same direction as the current flow does not diminish it, in both cases we found that small angle scattering does not contribute to the conductivity. In particular for the Boltzmann equation this was included through the transport scattering time τ_{tr} which we found to be twice as large as the scattering time τ . In the case of linear response this was included only when we added the contribution of vertex corrections to the current-current correlation function. This shows the vital importance of vertex corrections which should be assumed important unless one can show otherwise [66].

4.3 Tuning the Fermi-energy in real materials

We now switch to a brief discussion of experimental concerns in relation to transport on 3d topological insulator surfaces. To measure the transport properties of a topological insulator's surface it is important to isolate the surface conductivity from any contributions due to the bulk. In a theoretical calculation it is a simple task to ignore the bulk and consider a perfectly insulating TI where only conduction channels come from the surface states. In reality the isolation of the surface of a real material is far from trivial. Additionally, once the surface has been isolated, it is often desirable to measure quantities as a function of chemical potential. In this section we briefly discuss how both of these issues can be addressed experimentally.

4.3.1 Compensation of the bulk

We mentioned in chapter 1 that the archetypal 3d TI, Bi_2Se_3 is most naturally an n -type semi-conductor with a slight electron doping of its bulk [16]. Its cousin, Bi_2Te_3 , is generally p -type. Despite the multitude of topological “insulators” with a single Dirac-like surface or edge state, the vast majority are actually not very good insulators with some kind of bulk doping. As such in experimental samples a large bulk conductivity often completely obscures any contribution from the surface.

In order to fully access a TI’s surface states and associated physics it is therefore desirable to suppress the bulk occupation by tuning the Fermi energy deep into the bulk band gap. Luckily semi-conductor physics has taught us various methods to achieve this [83]. Here we focus on the method relevant to our experiments in the following chapters, namely, the introduction of charge dopants [16]. These dopants can either accept or donate electrons from or to ions within the material. Adding the right composition of acceptors and/or donors to a material can lead to an almost perfect compensation and very little occupation of the bulk bands.⁸

Perhaps the most prominent example of a compensated TI that achieves a high bulk resistivity is $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ ($y \geq 1$) [54]. By varying the composition of Sb – which provides donors – and Se – which provides acceptors – it is possible to achieve bulk resistivities of the order of several Ωcm (see Fig. 4.2a), with very good isolation of the surface contribution at low temperatures. In $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_{3-y}\text{Se}_y$ compensation has a second advantage: The position of the Dirac point change its location within the bulk band gap when dopants are added, as shown in Fig. 4.2b. This enables the Dirac point to be “engineered” to lie close to the Fermi-level and deep in the bulk gap[82].

4.3.2 Surface gating

Although doping is a sufficient process to ensure that the bulk chemical potential lies close to the TI’s Dirac point, it does not allow for a fine-tuned adjustment of the surface Fermi energy. Furthermore, doping varies the chemical composition of a sample which is difficult to change after fabrication and, even if possible, can change impurity concentrations making transport properties vary from sample to sample [16].

Instead, because a surface is only two dimensional, it is possible to achieve a large change in chemical potential through electrostatic gating [86]. A gate works by applying a voltage across the surface and through this shifting the position of the chemical potential. In general there are two contributions to the change in the surface’s energy when a gate voltage is applied: 1) The direct contribution to the change in surface energy $\Delta\mu$ and 2) The change resulting from the electron-static potential of the surface’s electron due to their (geometric) capacitance C_G . The gate voltage required to induce a charge Q on a

⁸Although desirable to increase the ratio of surface to bulk conductivity, adding these donors is not totally without consequence. Even at full compensation the bulk can contain charge puddles centered around these dopants[77, 78, 84]. These can even lead to novel resistivity properties of the bulk [85], but are not of relevance for our current discussion.

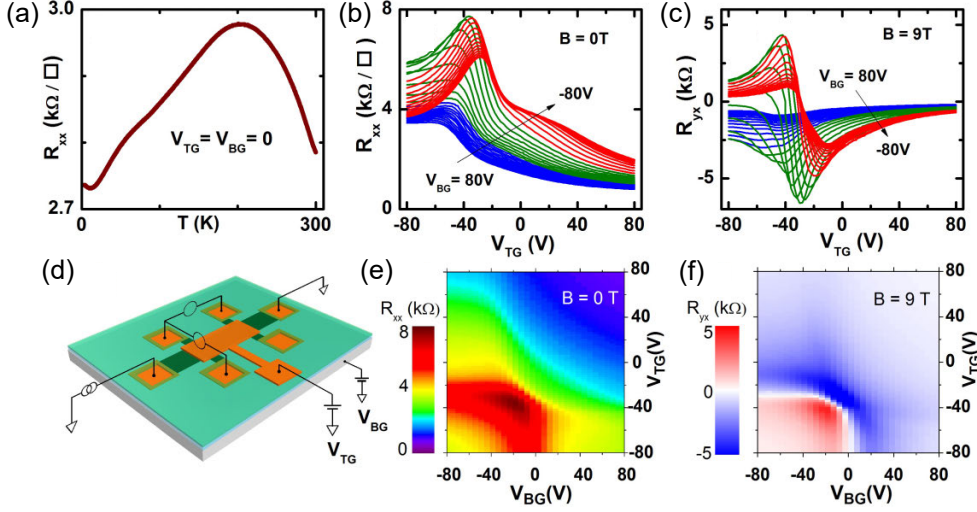


Fig. 4.2 **Dual surface gating of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$** : . (a) Temperature dependence of resistivity in a 17-nm-thick $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ sample at zero gate voltages ($V_{\text{TG}} = V_{\text{BG}} = 0$). (b), (c) Gate-voltage dependencies of longitudinal and hall resistivity at 2 K. (d) Schematics of the dual-gate Hall-bar device and the measurement configuration. (e), (f) Gate-voltage dependencies of the Hall conductivity in a perpendicular magnetic field of 9 T at 2 K. (Reproduced from Publication 1 [70].)

TI's surface is

$$V_G = \frac{e^2 \Delta n(\mu)}{C_G} - \frac{\Delta \mu}{e} = \frac{Q}{C_G} + \frac{Q}{C_Q}, \quad (4.27)$$

where $n(\mu)$ is the electron density, C_G is the geometric capacitance of the device, and C_Q the “quantum capacitance” [87]. In almost all experimental situations the energy due to the geometric capacitance is several orders of magnitude larger than the chemical potential, such that it is the dominant term and $V_G \approx \frac{1}{C_G} e n(\mu)$. The exception is very close to the Dirac point where the electron density, $n(\mu) \approx \frac{\mu^2}{4\pi v_F^2 \hbar^2}$, and the associated electrostatic energy are both small, meaning that the direct shift of the chemical potential μ by the gate can become important.

A single gate on top of a surface makes it possible to tune the Fermi-energy through the Dirac point of that surface. The position of the chemical potential on other surfaces, however, can still be far from the Dirac point and provide a short-circuiting mechanism where current can flow between contacts via that surface. It is therefore most desirable to have a *dual gate* device which can simultaneously adjust the chemical potential on the top and bottom surface of the topological insulator [88]. The resistance traces of a $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ device with surface dominated transport – the same used in the discussion in the next chapter – is shown in Fig. 4.2, we see that by tuning both gates simultaneously we are able to ascertain the location of the maximum in resistance since the dual gate provides great control of the chemical potential of the whole device [70].

Summary and outlook

We have seen how the conductivity of a disordered topological insulator surface can be calculated both semi-classically and using linear response theory. In particular we saw the importance that the angular factor $1 - \cos \theta$ in transport scattering time τ_{tr} and vertex corrections within linear response theory, implement the fact that forward or small angle scattering have little impact on the decay of the current. Finally we discussed some experimental challenges in measuring surface transport from a topological insulator surface. In particular we saw how the bulk contribution to conductivity can be reduced using the concept of compensation and how it is possible to tune the Fermi-energy of a material by applying a gate and inducing surface charge.

Having established the method to calculate conductivities of a TI surface and how experiments can be performed, in the next 2 chapters we will implement these concepts in two different scenarios: In chapter 5 we will see that it is possible to lift the topological protection of the surface states by applying an in-plane magnetic field. This results in an anisotropic magnetoresistance which we will calculate theoretically and discuss experimental measurements of such an effect. In chapter 6 we will see that the finite geometry of a topological insulator nanowire results in a quantisation of the angular momentum around the wire. We will show that this quantisation leads to a unique and experimentally measurable transport fingerprint as a function of chemical potential.

Planar Hall effect from the surface of topological insulators

We have seen throughout the previous chapters that the surface states of topological insulators are described by a massless two-dimensional Dirac Hamiltonian and, as a result, are topologically protected. In particular, we saw in chapter 2 the electron spin-momentum locking forbade back-scattering from non-magnetic scatterers. In chapter 3 we discussed a direct consequence of this: The faster $1/r^2$ decay of LDOS oscillations around an impurity. In chapter 4 the protection from back-scattering lead to the unique transport properties of a TI surface, resulting in a conductivity 4 times higher than in the absence of the protection. In this chapter we describe theoretically and experimentally how it is possible to lift this topological protection in a controlled manner whilst preserving the Dirac physics of the TI surface. This is achieved by the application of a magnetic field in the plane of the TI surface being measured. As a result the magnetoresistance measured in the directions parallel and perpendicular to this in-plane field is anisotropic and, associated with this **anisotropic magnetoresistance (AMR)**, is a **planar Hall effect (PHE)**.

The layout of this chapter is as follows: To understand the origin of anisotropic scattering on a TI surface we first discuss the simplest mechanisms for lifting a TI's topological protection and how it is possible to lift this protection in a controlled manner by an in-plane field. In the majority of this chapter we establish a theory for the resulting anisotropic magnetoresistance based on a self-consistent T-matrix approximation and Kubo formula, which were introduced in previous chapters. Having established the theoretical basis for such an effect we will see that such an affect is found in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ thin films. We will discuss the experimental signatures of this effect and what we can learn from the experimental results through a detailed comparison of experiment and theory. In particular, we will show that both theory and experiment have a characteristic “two peak” structure of AMR as a function of chemical potential and that this can be utilised to ascertain several pieces of information about the characteristics of the TI surface.

This chapter is based on the publication 1: “Planar Hall effect from the surface of topological insulators” Nat. Commun. 8, 1340 (2017) and its supplementary material [70].

5.1 Lifting topological protection

Topological protection can be lifted in a variety of scenarios. For example, in a very thin TI film a significant overlap of the wavefunctions of the top and bottom surface states can

Chapter 5. Planar Hall effect from the surface of topological insulators

occur as they penetrate into the bulk of the TI [89]. This overlap results in a Hamiltonian of the form

$$H_{\text{hyb}}^{\text{el}} = \sum_{\mathbf{k}} \psi_{\mathbf{k}}^{\dagger} [v_F \tau_z (k_x \sigma_y - k_y \sigma_x) + \Delta \tau_x] \psi_{\mathbf{k}}, \quad (5.1)$$

where $\tau_z = \pm 1$ indicates the surface, such that opposite surfaces have opposite chirality; the hybridisation Δ is a measure of the matrix element between the top and bottom surface, resulting in off-diagonal elements in the 4×4 matrix space defined by the τ and σ matrix products; and $\psi_{\mathbf{k}}$ are our usual electron spinors, here acting as creation and annihilation operators. For small Δ the corresponding wave functions are localised about one of the two surfaces and on each surface these states have energy eigenvalues $E_{\text{hyb}} = \pm \sqrt{v_F^2 (k_x^2 + k_y^2) + \Delta^2}$. As such, in principle any surface hybridisation opens up a gap at the Dirac point, leading to a loss of topological protection and therefore enabling back-scattering [90]. In reality however the surface states' wave functions decay exponentially into the TI bulk and so only the thinnest of TI films – less than ~ 10 nm thick for Bi_2Se_3 – will have a hybridisation large enough to have any noticable experimental consequences. For the 17 nm TI films used in the experiment considered in this chapter, surface hybridisation will be much smaller than any other relevant energy scale and therefore can be safely neglected.

When considering a single isolated TI surface, as discussed in chapter 1, the surface states are protected by time-reversal symmetry (TRS). Hence by breaking TRS the topological protection will be lifted. The most natural way to do this is by applying a magnetic field. To begin with we consider a field $\mathbf{B} = B\hat{z}$ perpendicular to the TI surface. The resulting Hamiltonian is

$$H_{B_{\perp}}^{\text{el}} = \sum_{\mathbf{k}} \psi_{\mathbf{k}}^{\dagger} [v_F ((k_x - A_x) \sigma_y - (k_y - A_y) \sigma_x) + g\mu_B B \sigma_z] \psi_{\mathbf{k}}, \quad (5.2)$$

where $\mathbf{A} = (A_x, A_y, 0)$ is the vector potential of the magnetic field and g the g-factor of the TI surface due to the Zeeman effect. We immediately see that a perpendicular field, whilst breaking TRS, also changes the Dirac physics of the electrons on the surface in two important ways: 1) The Zeeman term – which can be unusually large in TIs due to their large g-factors – acts as a mass of size $g\mu_B B$ in the Dirac Hamiltonian. This means that a gap is opened in the surface states. As we saw in chapter 2, electron states close to the bottom of the conduction band or hole states close to the top of the valence band of this gapped Dirac spectrum will have poor spin-momentum locking. 2) Even more significantly, the orbital terms due to the vector potential \mathbf{A} cause the formation of Landau levels and results in the well known Hall physics of electrons in a magnetic field (we discuss this in 3d Dirac materials in chapter 7). Whilst breaking TRS with a perpendicular magnetic field can result in extremely interesting quantum Hall insulator physics, even small fields completely change the Dirac physics of the surface and so are a poor method for assessing the lifting of the surface's topological protection in a controlled manner.

5.1. Lifting topological protection

Although it is not possible to retain the Dirac physics with a perpendicular field, any magnetic field will break TRS, including a field applied parallel to the TI surface. Since the vector potential for a parallel magnetic field can be written in the form $\mathbf{A} = (0, 0, A_z)$, the Hamiltonian for the surface states in the presence of a parallel magnetic field – for example in the x -direction i.e. $\mathbf{B} = B\hat{x}$ – reads

$$H_{B\parallel}^{\text{el}} = \sum_{\mathbf{k}, \alpha, \beta} \psi_{\mathbf{k}}^{\dagger} \left[v_F (k_x \sigma_y - \left(k_y - \frac{g\mu_B B}{v_F} \right) \sigma_x - \mu \mathbb{1}) \right] \psi_{\mathbf{k}}, \quad (5.3)$$

where we have included a chemical potential μ . Therefore a parallel field will not affect the surface states apart from a shift of the Dirac cone in the momentum space $k_y \rightarrow \tilde{k}_y = k_y - g\mu_B B/v_F$ proportional to the Zeeman coupling $g\mu_B B$. In particular the shift in momentum space can be gauged away by an additional phase of the electron operators $\psi(\mathbf{r}) \rightarrow e^{ig\mu_B B y/v_F} \psi(\mathbf{r})$ and the Dirac physics of the surface states will be unaffected [70].¹

In reality, however, any surface will contain some sort of disorder. The uniform magnetic field parallel to the surface, whilst having no direct effect on the surface-states themselves, will have an effect in the vicinity of this disorder. In particular, the impurities themselves will have a different g -factor, g^{imp} , to the surrounding surface. Therefore, the impurities generate local magnetic fields at random positions on the surface that are parallel or anti-parallel to the magnetic field.² The basic idea explored in this chapter is that these local magnetic fields can induce spin-flip scattering when the spin of the electron is not aligned with them, in other words, momenta parallel to the applied magnetic field are no longer protected from backscattering, whereas momenta perpendicular to the field remain protected. We see that such a scenario fulfils our quest to find a way to lift the protection of the surface Dirac fermions without a complete breakdown of the Dirac physics.

To describe the effect of these local magnetic fields, we need to add the magnetised impurities, located at random positions \mathbf{R}_i , to the Hamiltonian of the surface states, giving

$$H_{B\parallel} = H_{B\parallel}^{\text{el}} + \sum_{\alpha, \beta} ((\epsilon - \mu) \delta_{\alpha\beta} - g^{\text{imp}} \mu_B \sigma_x) d_{\alpha}^{\dagger} d_{\beta} + \sum_{\mathbf{k}, \alpha, i} V e^{-i\mathbf{k}\cdot\mathbf{R}_i} \psi_{\alpha\mathbf{k}}^{\dagger} d_{\alpha} + \text{h.c.}, \quad (5.4)$$

where the first term is simply the surface state Hamiltonian Eq. (5.3), the second term describes a localized resonance with energy ϵ which, in the third term, hybridises with the continuum states with hybridisation strength V . We have taken the magnetic field to be in the x -direction. This model is not expected to give a microscopic description of the actual disorder in the experiment we discuss at the end of the chapter. Nonetheless, it will provide a minimal model capable of capturing the field-induced anisotropic back-scattering

¹One caveat to this statement is that, since time-reversal symmetry is broken, *in principle* a gapless surface state is only guaranteed in high-symmetry directions where further symmetries protect the surface states. For instance, if we consider the (111) surface of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ from the experiment considered in this chapter – which has an $R\bar{3}m$ symmetry – we are only guaranteed a gapless surface state for directions where a mirror symmetry exists, this means a field applied in the $[1\bar{1}0]$ direction and equivalent directions rotated 60° . In practice, however, away from these high symmetry directions any gap will be extremely small.

²The direction of the random fields will depend on the size of the g^{imp} relative to the electron g -factor but ultimately this direction does not impact the opening of a back-scattering channel discussed later.

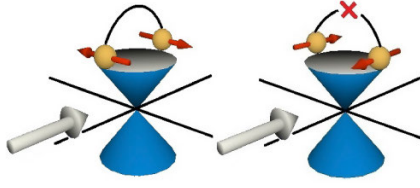


Fig. 5.1 **Field induced Spin-flip scattering:** Schematic of spin-flip scattering. Left: When the field and electron momentum are parallel or anti-parallel (spin, red arrow, is perpendicular) magnetised impurities allow spin-flip scattering. Right: When the field is perpendicular to the momentum (parallel or anti-parallel to the spin) back-scattering remains forbidden.

and the interplay of magnetic (spin-flip) and non-magnetic (non-spin-flip) scattering which allows us to model the lifting of the topological protection.

Of particular importance is that, in the presence of a magnetic field, the impurities are no longer diagonal in spin-space, but possess a magnetic component due to their magnetisation along the direction of the in-plane field. Therefore a new scattering channel is opened up with matrix elements

$$M_{\mathbf{k},\mathbf{k}'}^B = (g^{\text{imp}} \mu_B B)^2 |\langle \mathbf{k} | \sigma_x | \mathbf{k}' \rangle|^2 = (g^{\text{imp}} \mu_B B)^2 \sin^2 \left(\frac{\theta_{\mathbf{k}} + \theta_{\mathbf{k}'}}{2} \right), \quad (5.5)$$

where we use the same notation as in chapter 2. This matrix element is highly anisotropic: For momenta parallel (anti-parallel) to the magnetic field – here $\theta_{\mathbf{k}} = 0^\circ$ ($\theta_{\mathbf{k}} = 180^\circ$) – back-scattering is the dominant scattering direction due to the magnetic component of the impurity σ_x . In contrast, for momenta perpendicular to the field – here $\theta_{\mathbf{k}} = \pm 90^\circ$ – the matrix element for backscattering is identically zero and so backscattering remains forbidden in this direction. It is this anisotropy in scattering that is the basis for the AMR.

5.2 Scattering from magnetised resonances

As in previous chapters we will use the T-matrix formalism to model scattering from disorder. Our surface states will once again be described by the full disorder averaged Green's function $\mathcal{G}(\mathbf{k}, \mu)$, with a self-energy discussed below. In previous chapters we considered impurities that were structureless local potentials $\mathbf{U}(\mathbf{r} - \mathbf{r}_i)$. In contrast the local resonances in the Hamiltonian Eq. (5.4) enable electrons to live on the resonance and these states on the resonance are then effectively scattered by the continuum of surface states. We can therefore associate with the resonance its own (free) Green's function³

$$g_0^{\text{res}}(\mu) = (\Omega + \mu - \epsilon - g\mu_B[\mathbf{B} \cdot \boldsymbol{\sigma}])_{\alpha\beta}^{-1}, \quad (5.6)$$

which is a 2×2 matrix.

We start by considering the clean surface states, described by the free Green's function $G_0(\mathbf{k}, \mu)$ (see chapter 2), scattering off a single resonance [62]. This scattering can be

³We will keep a finite frequency Ω throughout. The reason why will become apparent later.

5.2. Scattering from magnetised resonances

represented diagrammatically – after impurity averaging over the impurity’s position – as

$$\begin{aligned}
\Rightarrow \text{double line with arrow } k &= \text{single line with arrow } k + \text{single line with arrow } k \text{ and two } V \text{ vertices connected by a dashed line} + \text{single line with arrow } k \text{ and three } V \text{ vertices connected by two dashed lines} + \dots \\
&= \text{single line with arrow } k + \text{single line with arrow } k \text{ and two } V \text{ vertices connected by a dotted line}, \quad (5.7)
\end{aligned}$$

where dashed lines are the resonance Green's function and solid (double lines) the surface state Green's function. In the second line we have defined the full Green's function of resonance states being scattered by the surface states as

$$\begin{aligned}
\text{-----} &= \text{-----} + \text{-----} \overset{V}{\bullet} \xrightarrow{k_1} \overset{V}{\bullet} \text{-----} + \text{-----} \overset{V}{\bullet} \xrightarrow{k_1} \overset{V}{\bullet} \text{-----} \overset{V}{\bullet} \xrightarrow{k_2} \overset{V}{\bullet} \text{-----} + \dots \\
&= \text{-----} + \text{-----} \overset{V}{\bullet} \xrightarrow{k} \overset{V}{\bullet} \text{-----}, \tag{5.8}
\end{aligned}$$

where we integrate over each \mathbf{k}_i of the surface state Green's functions.

We see that this diagrammatic expansion is a geometric series, such that summing over it gives the full Green's function of the resonance

$$\text{.....} = g^{\text{res}}(\mu) = g_0^{\text{res}}(\mu)(1 + \Delta(\mu)g_0^{\text{res}}(\mu) + (\Delta(\mu)g_0^{\text{res}}(\mu))^2 + \dots) \quad (5.9)$$

$$= (\Omega + \mu - \epsilon - \Delta(\mu) + g\mu_B [\mathbf{B} \cdot \boldsymbol{\sigma}])^{-1}. \quad (5.10)$$

Here we have defined the resonance Green’s function self-energy $\Delta(\mu)$ – which we call the “hybridisation function” –, it comes from the resonance states scattering from the continuum surface states, such that

$$\Delta(\mu) = V^2 \int \frac{d\mathbf{k}}{(2\pi)^2} G_0(\mathbf{k}, \mu), \quad (5.11)$$

and is 2×2 matrix.

Returning to the expansion of the surface states Green's function in Eq. (5.7) we recognise that the T-matrix for the surface states is $T(\mu) = V^2 g^{\text{res}}(\mu)$, where the factor V^2 comes from the fact we must hop on and hop off the resonance picking up a factor V in each direction. In particular the T-matrix is diagonal in spin space at zero-field but develops an off-diagonal component when the field is turned on. This off-diagonal component enables the spin-flip scattering discussed above which will be fleshed out below.

5.2.1 Surface state self-energy

We now consider a finite density of resonances n_{imp} . We have seen that, after impurity averaging, the non-crossing approximation for dilute impurities leads to a surface state self-energy

$$\Sigma(\mu) = n_{\text{imp}} T_{kk}(\omega) = n_{\text{imp}} V^2 g^{\text{res}}(\mu), \quad (5.12)$$

Chapter 5. Planar Hall effect from the surface of topological insulators

in terms of the full resonance Green's function as outlined above. As in chapter 2, we can further increase the number of diagrams included in this definition by making it self-consistent. We do this by replacing the free Green's function $G_0(\mathbf{k}, \omega)$ in the definition of the hybridisation function with the full disorder averaged Green's function, such that

$$\Delta(\mu) = V^2 \int \frac{d\mathbf{k}}{(2\pi)^2} \mathcal{G}(\mathbf{k}, \mu). \quad (5.13)$$

Let us first consider the zero-field self-energy, setting $B = 0$. In this case, as in chapter 2, the angular integral over the off-diagonal components of the Dirac Green's function cancels in the self-consistent equation Eq. (5.12), and the off-diagonal components of self-energy Σ_{12} are zero. The remaining diagonal elements satisfy $\Sigma = \Sigma_{11} = \Sigma_{22}$, such that

$$\Sigma_0(\mu) = n_{\text{imp}} |V|^2 \left(\Omega + \mu - \epsilon + \frac{|V|^2}{4\pi v_F^2} (\omega + \mu - \Sigma(\mu)) \ln \left(\frac{-\Lambda^2}{(\Omega + \mu - \Sigma(\mu))^2} \right) \right)^{-1}, \quad (5.14)$$

which is similar to what we have seen in chapter 2. Three examples of the real and imaginary part of the zero-frequency self-energy $\Sigma_{11}(\mu)$ are shown in the top row of Fig. 5.2, these are found by numerically solving Eq. (5.12). At zero frequency, $\Omega = 0$, the equation has two distinct regimes: (i) A metallic regime, when $|\mu|$ is large. Here both real and imaginary parts of self-energy are small, $\Sigma \sim 1/\mu$. (ii) An impurity dominated regime near the Dirac point, where the self-consistency becomes important.

The on-set of the impurity dominated regime is characterised by a large increase in the absolute value of the imaginary part of the self-energy; correspondingly, as a result of Kramers-Kronig relation, there are two maxima in the real part of the self-energy. For the particle-hole symmetric resonances, $\epsilon = 0$, the self-energy is purely imaginary at the Dirac point, $\mu = 0$. The energy scale here, $\Sigma(\mu = 0) = -i\Gamma_0$, defines the width of the impurity dominated regime. From Eq. (5.14) we see that Γ_0 is given by self-consistently solving

$$\Gamma_0 = \sqrt{\frac{2\pi n_{\text{imp}} v_F^2}{\ln \left(\frac{\Lambda}{\Gamma_0} \right)}}. \quad (5.15)$$

This energy scale sets many of the features we will see in the coming sections.

Finally let us distinguish between the strong and weak scattering regimes. In chapter 2 this was set by the strength of the potential fluctuations u_0 . For the resonances currently under consideration strong impurity scattering is realised for *small* ϵ , that is, when scattering is approximately resonant. An inspection of the denominator in Eq. (5.14) reveals that strong, approximately resonant impurity scattering is realised when $|\epsilon| \lesssim V^2 \Gamma_0^2 / (4\pi v_F^2) = \eta_0$. In contrast, highly off-resonance impurities, $\epsilon \gg \eta_0$, are weak. This means that scattering from resonances with an energy close to the Dirac point is *always* effectively strong, only when the resonance energy ϵ is substantially above or below the energy of the Dirac point are we in a weak scattering regime [65, 91]. This emphasises the importance of using the full T-matrix in the calculating the self-energy.

5.2. Scattering from magnetised resonances

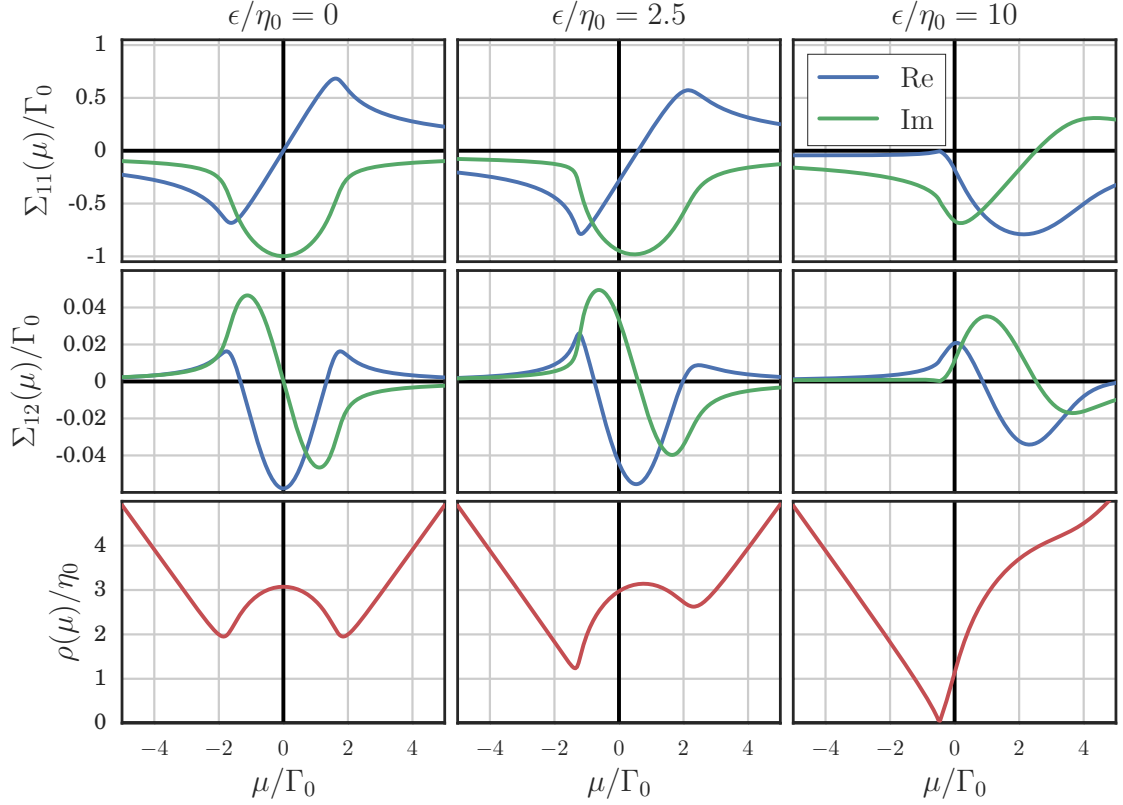


Fig. 5.2 **Self energy and density of states due to resonant impurities:** Top row: The diagonal component of self-energy $\Sigma_{11}(\mu)$ for $\epsilon/\eta_0 = 0, 2.5, 10$ respectively. Middle row: The off-diagonal component of self-energy $\Sigma_{12}(\mu) \sim (\Sigma_{11}(\mu))^2$ for $B = 0.25\eta_0$ and the same ϵ parameters. Bottom row: density of states. Impurity scattering dominates close to the Dirac point which leads to distinct features in all quantities. Parameters: $\Lambda = 10$, $V = \sqrt{4\pi}$.

5.2.2 Density of states

Before we turn the magnetic field on, let us consider the density of states. As in chapter 2, this is given by the trace of the Dirac Green's function,

$$\rho(\mu) = -\frac{1}{\pi} \text{Im} \left\{ \text{Tr} \int \frac{d^2k}{(2\pi)^2} \mathcal{G}(\mathbf{k}, \mu) \Big|_{\Omega=0} \right\} = \text{Im} \left\{ \frac{2n_{\text{imp}} V^2}{\pi \Sigma_0(\mu)} \right\}, \quad (5.16)$$

where the second line can be obtained by inserting the self-consistent equation Eq. (5.14) solved for the hybridization function. The density of states is shown in the final row of Fig. 5.2. From this we see that in the metallic regime the density of states is linear, as in the clean system. However, similar to strong scattering discussed in chapter 2, close to the Dirac point the density of states is strongly affected by the presence of resonances, with new states created between the bounds of the regime set by the energy scale Γ_0 . The transition from the impurity dominated regime to the metallic regime is characterised by the presence of a minima in the density of states at the location of the peaks in $\text{Re}\Sigma_{11}$.

5.2.3 Finite magnetic field

At finite \mathbf{B} – which we take in the following discussion to be in the x -direction – the self-consistent T-matrix equation, Eq. (5.12), becomes a full matrix equation with non-zero off-diagonals of the self-energy matrix $\Sigma(\mu)$ and hybridization function $\Delta(\mu)$. The surface state's self-energy is now

$$\begin{aligned}\Sigma(\mu) &= n_{\text{imp}} V^2 \begin{pmatrix} \Omega + \mu - \Delta_{11}(\mu) - \epsilon & g\mu_B B - \Delta_{12}(\mu) \\ g\mu_B B - \Delta_{12}(\mu) & \Omega + \mu - \Delta_{11}(\mu) - \epsilon \end{pmatrix}^{-1} \\ &\approx \Sigma_0(\mu) \mathbb{1} + g\mu_B B \frac{\Sigma_0(\mu)^2}{n_{\text{imp}} V^2} \sigma_x,\end{aligned}\tag{5.17}$$

where the last line is valid for small \mathbf{B} and $\Sigma_0(\mu)$ is the zero-field self-energy (above). Examples of the real and imaginary part off-diagonals of the self-energy, $\Sigma_{12}(\mu)$, calculated numerically at a finite field are shown in the middle row of Fig. 5.2.

In chapter 2 we discussed that the physical interpretation of $\text{Im}\Sigma_{11}(\mu)$ is related to the scattering time $1/\tau$, including the surface state's protection from back-scattering. Similarly the off-diagonal of the self-energy, in particular $\text{Im}\Sigma_{12}(\mu)$, is a measure of the *spin-flip* scattering rate $1/\tau_s$, this is the rate of scattering due to the magnetised component of the resonances – we saw above such scattering is *not* protected from back-scattering when the spin is perpendicular to the field. The ratio of spin-flip to total scattering rate is therefore $\frac{1/\tau_s}{1/\tau} = \frac{\text{Im}\Sigma_{12}(\mu)}{\text{Im}\Sigma_{11}(\mu)}$. Importantly we find from the self-energy in a small finite field, Eq. (5.17), that $\text{Im}\Sigma_{12}(\mu)$ is proportional to $\text{Im}[\Sigma_{11}(\mu)^2] = 2\text{Im}\Sigma_{11}\text{Re}\Sigma_{11}$. As a result the peaks we find in $\text{Re}\Sigma_{11}$ that occur at the onset of the impurity dominated regime in Fig. 5.2 also lead to peaks in spin-flip scattering rate.

The maxima in spin-flip scattering will be substantial for scatterers close to resonance due to the large impurity dominated regime in the strong scattering limit. On the other hand, in the weak scattering regime, $\epsilon \gg \eta_0$, the peaks in the spin-flip scattering rate $\text{Im}\Sigma_{12}(\mu)$ will vanish since the impurity dominated regime is exponentially suppressed in the Born limit, as we also saw in chapter 2.

5.3 Conductivity in a parallel field

We now have all the components required to calculate the conductivity resulting from these magnetised resonances. The calculation is largely the same as we have performed in the previous chapter, the crucial physical difference is that the self-energy now contains off-diagonal components.

Whilst we will include vertex corrections later in our calculation, it is instructive to first ignore vertex corrections and consider the bare bubble. We saw in the last chapter that this is given by

$$\sigma_{\alpha\alpha}^{(0)} = \frac{e^2 v_F^2}{2\pi} \text{Re} \int \frac{d^2 \mathbf{k}}{(2\pi)^2} \text{Tr} \left\{ \sigma_\alpha \mathcal{G}(\mathbf{k}, \mu) \sigma_\alpha \mathcal{G}^\dagger(\mathbf{k}, \mu) \right\},\tag{5.18}$$

5.3. Conductivity in a parallel field

where the current operator satisfied $\hat{j}_\alpha = ev_F(\sigma_y, -\sigma_x)_\alpha$. As in the previous sections we will, without any loss of generality, take the magnetic field to lie in the x -direction such that $\sigma_{xx} = \sigma^\parallel$ is the conductivity parallel to the field and $\sigma_{yy} = \sigma^\perp$ is the conductivity perpendicular to the field.

From Eq. (4.20) in the last chapter the conductivity parallel and perpendicular to the field (taking $\Omega \rightarrow 0$) is

$$\begin{aligned} \sigma^{\parallel/\perp} &= \frac{e^2 v_F^2}{2\pi} \int \frac{d^2 \mathbf{k}}{(2\pi)^2} \frac{|\mu - \Sigma_{11}|^2 \pm (v_F k_\parallel^2 - |v_F k_\perp - \Sigma_{12}|^2)}{[(\mu - \Sigma_{11})^2 - (v_F^2 k_\parallel^2 + (v_F k_\perp - \Sigma_{12})^2)]^2} \\ &\approx \sigma_0 \left(1 \mp \frac{c}{2} \left(\frac{\text{Im}\Sigma_{12}}{\text{Im}\Sigma_{11}} \right)^2 \right), \end{aligned} \quad (5.19)$$

where Σ_{11} and Σ_{12} are the diagonal and off-diagonal contributions to the self energy from the last section. Numerically we find the prefactor c to be $0.8 - 1.2$, depending on parameters. The second equality was derived by using $\Sigma_{12} \ll \Sigma_{11}$ and absorbing the real part of Σ_{12} by a shift of k_y .

We now discuss the main result of this chapter: The anisotropy of magnetoresistance induced by the magnetised impurities. Experimentally it is easiest to measure the resistivity $\rho^{\parallel/\perp} = \frac{1}{\sigma^{\parallel/\perp}}$ and so we will also discuss our theoretical results in terms of resistivities. From equation Eq. (5.19) we see that the resistivity is *higher* when the field is parallel to the current and *lower* when the field is perpendicular. In particular, the difference in resistivity is

$$\Delta\rho(\mu) = \rho^\parallel - \rho^\perp \approx c \left(\frac{\text{Im}\Sigma_{12}}{\text{Im}\Sigma_{11}} \right)^2, \quad (5.20)$$

where to obtain the last approximation we used that $\frac{\text{Im}\Sigma_{12}}{\text{Im}\Sigma_{11}} \ll 1$. We see that this difference is proportional to the ratio of spin-flip and non spin-flip scattering discussed in the previous section. The effect is quadratic in B for small B as $\text{Im}\Sigma_{12}$, the spin-flip scattering rate, is linear in B for small B .

The ratio of the difference in resistivity $\Delta\rho(\mu)$ and total resistivity, using the full self-energy calculated self-consistently, is plotted in Fig. 5.3 (this plot also includes vertex corrections, see below. There is also an example of averaging over parameters in this plot, see Appendix C). A clear two-peak structure emerges in the spin-flip scattering rate due to the maxima in $\text{Im}\Sigma_{12}$. The asymmetry of the two peaks is controlled by the resonance energy ϵ . The location of the peaks track precisely the minima in the density of states also shown in Fig. 5.3, we will see in the next section that this double peak structure allows us to extract several key pieces of information from experimental measurements of this effect.

5.3.1 Vertex corrections

There is no reason to assume that vertex corrections can be neglected in this case, especially given the fact our scatterers are so highly anisotropic. As in the last chapter we can calculate the vertex correction by mapping each pair of Green's functions and T-matrices to the

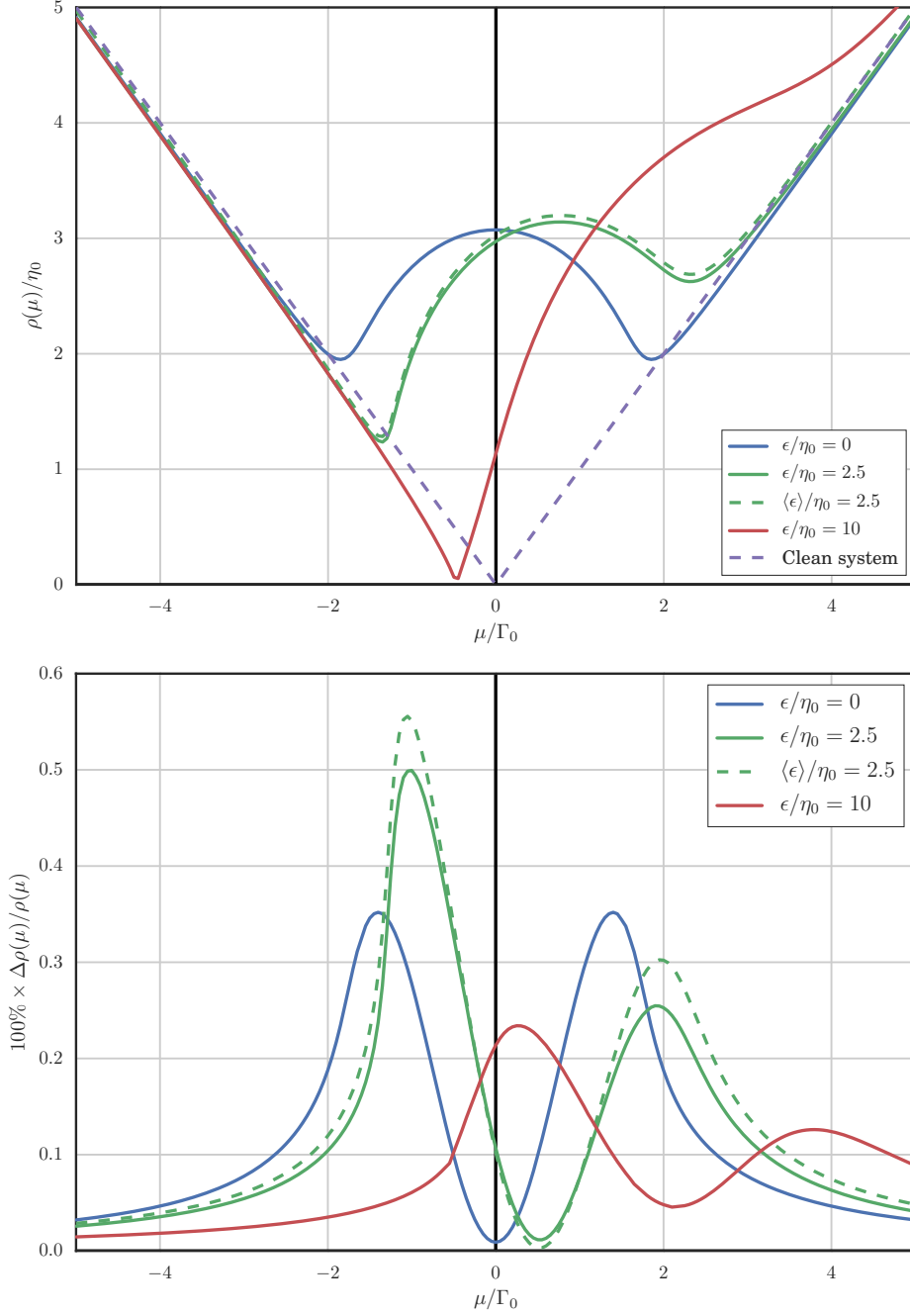


Fig. 5.3 **Peaks in AMR as a function of chemical potential:** Top: Density of states for various parameters (see Fig. 5.2 for self-energy, dashed lines are averaged quantities, see Appendix C). Bottom: Size of anisotropy in magnetoresistance in percent. The locations of peaks in the AMR correspond to minima in the density of states leading to a clear two-peak structure. This can be traced back to peaks in $\text{Im}\Sigma_{12}$, the spin-flip scattering rate. (Parameters for solid lines: $\Lambda = 10$, $V = \sqrt{4\pi}$. See Appendix C for details of averaged parameters, dashed lines.)

5.3. Conductivity in a parallel field

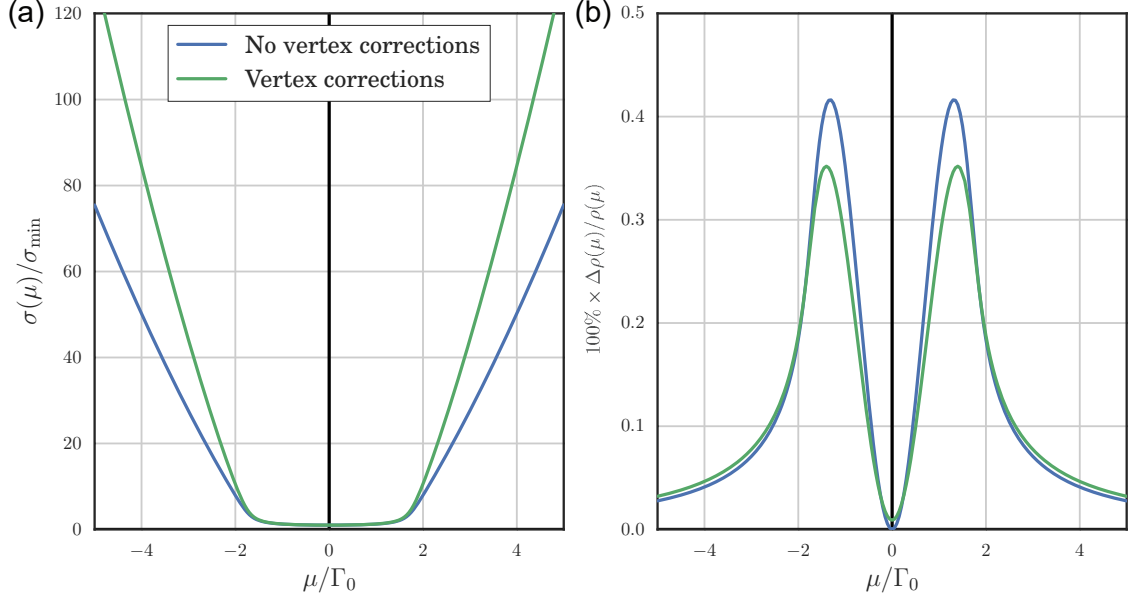


Fig. 5.4 **Vertex corrections:** (a) The conductivity and (b) dimensionless resistivity anisotropy with and without vertex corrections for the particle hole symmetric system $\epsilon = 0$. Vertex corrections approximately double the conductivity in the metallic regime but have only a small effect near the peaks of the resistivity anisotropy.

4×4 Kronecker products

$$\mathcal{M}(\mu) = \int \frac{d^2\mathbf{k}}{(2\pi)^2} \mathcal{G}_\Omega(\mathbf{k}, \mu) \otimes \mathcal{G}^\dagger(\mathbf{k}, \mu) \quad \text{and} \quad \mathcal{T}(\mu) = T_\Omega(\mu) \otimes T^\dagger(\mu), \quad (5.21)$$

where the subscript Ω indicates that there is a frequency difference of Ω between the first and second components of this product. As previously, the full conductivity is

$$\sigma_{\alpha\beta}(\mu) = \lim_{\Omega \rightarrow 0} \text{Re} \left\{ e^2 v_F^2 \int \frac{d\omega}{\pi} \frac{\Delta n_F(\Omega)}{\Omega} \sigma^\alpha (\mathcal{M}(\mathbb{1} - n_{\text{imp}} \mathcal{T} \mathcal{M})^{-1}) \sigma^\beta \right\}, \quad (5.22)$$

where $\Delta n_F(\Omega) = n_F(\Omega) - n_F(\Omega = 0)$. We have explicitly left the limit $\Omega \rightarrow 0$ in our formula since this must be taken only at the very end of the calculation. If we were to take the limit $\Omega \rightarrow 0$ *before* inverting the correction factor $(\mathbb{1} - n_{\text{imp}} \mathcal{T} \mathcal{M})$, then we would find that this matrix is singular at any finite magnetic field. The underlying reason is that the current operator, $\hat{j}_\alpha = \sigma_\alpha$, is of the same form as any magnetisation of the surface parallel to the field. As a result, a finite magnetisation leads to $\langle \hat{j}_\alpha \rangle \neq 0$ and a (non-dissipating) current flows perpendicular to the magnetic field, even for a vanishingly small electric field. This leads to a singularity in $(\mathbb{1} - n_{\text{imp}} \mathcal{T} \mathcal{M})$ since there are components of the matrix $\mathcal{M}(\mathbb{1} - n_{\text{imp}} \mathcal{T} \mathcal{M})^{-1}$ which describe a dissipationless flow and therefore must be infinite in the limit $\Omega \rightarrow 0$. If, however, we leave the limit $\Omega \rightarrow 0$ until the end of the calculation, since the contraction of this matrix with σ_α does not couple to these dissipationless modes, the true conductivity remains finite.

The effect of vertex corrections are shown in Fig. 5.4, the vertex corrections approximately double the conductivity in the metallic regime far from the Dirac point, but for the ratio $\Delta\rho/\rho$, as plotted in Fig. 5.3, they cause only a small reduction in the vicinity of the two maxima and no significant qualitative change in the AMR.

5.4 Anisotropic MR and planar Hall effect in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$

Now that we have demonstrated theoretically that an in-plane field is capable of lifting protection from back-scattering we turn to the observation of this effect on the surface of real topological insulators, namely in bulk-insulating $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ thin films. Experimentally several challenges must be overcome to obtain clear and clean data of any anisotropy of resistivity: (1) The topological insulator must actually be an insulator in its interior so that the surface transport is not entirely washed out by a bulk contribution. (2) To obtain data on how the effect changes with carrier density we must have control over the surface chemical potential. (3) We must be able to eliminate any angular dependence of the resistivity that comes from a misalignment of the magnetic field and topological insulator surface.

At the end of the last chapter we saw that compensation is capable of drastically reducing the bulk contribution, solving (1). We also saw that surface gating, in particular surface gating on both top and bottom surfaces, allows a fine tuning of the chemical potential and solves (2). In fact Fig. 4.2 from the previous chapter is the same device that the experiments discussed here are conducted on.

5.4.1 The planar Hall effect

It remains to find a way to eliminate any spurious contribution to the resistivity due to a misalignment of the sample and the applied magnetic field. A sample misalignment results in a small component of the magnetic field perpendicular to the TI surface. This out of plane component B_\perp will result in a standard orbital magnetoresistance (MR) ΔR_\perp^* . Although B_\perp is normally very small, we have also seen that any AMR is small – of the order of one percent of total resistance – and so the contribution from the orbital MR in a real experiment can be comparable to the amplitude of the AMR. To make matters worse, upon rotation of the magnetic-field, B_\perp , the projection of \mathbf{B} out of the plane, changes with a period of 180° . When combined with the standard $\Delta R_\perp^* \sim B_\perp^2$ behaviour of orbital MR for small B_\perp , a misalignment causes a $\sim \cos^2\varphi$ dependence, making it indistinguishable from the genuine AMR signal, which has the same dependence. As a result, the amplitudes of the AMR in actual experiments are not always reliable and may be heavily contaminated by the resistivity due to sample misalignment.

Luckily, however, there is a way to remove the contribution of a misalignment from experimental data. In a real experiment the current is measured in a fixed basis in the x - y plane of the TI surface. This is not, in general, the same basis as the resistivity anisotropy, which we have seen theoretically is diagonal in a basis along the $\hat{\mathbf{r}}_\parallel$ and $\hat{\mathbf{r}}_\perp$ directions,

5.4. Anisotropic MR and planar Hall effect in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$

parallel and perpendicular to the field. To calculate the actual resistivity measured in the fixed basis of an experiment we can simply need to perform a change of basis, such that the resistivity matrix becomes

$$\begin{aligned} \begin{pmatrix} E_x \\ E_y \end{pmatrix} &= \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \rho_{\parallel} & 0 \\ 0 & \rho_{\perp} \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} j_x \\ j_y \end{pmatrix}, \\ &= \begin{pmatrix} \rho_{\parallel} \cos^2 \varphi + \rho_{\perp} \sin^2 \varphi & (\rho_{\parallel} - \rho_{\perp}) \cos\varphi \sin\varphi \\ (\rho_{\parallel} - \rho_{\perp}) \cos\varphi \sin\varphi & \rho_{\parallel} \sin^2 \varphi + \rho_{\perp} \cos^2 \varphi \end{pmatrix} \begin{pmatrix} j_x \\ j_y \end{pmatrix}. \end{aligned} \quad (5.23)$$

Assuming a measurement configuration along the x -direction, such that $j_y = 0$, one obtains

$$\rho_{xx} = \frac{E_x}{j_x} = \rho_{\perp} + (\rho_{\parallel} - \rho_{\perp}) \cos^2 \varphi \quad \text{and} \quad \rho_{yx} = \frac{E_y}{j_x} = (\rho_{\parallel} - \rho_{\perp}) \cos \varphi \sin \varphi. \quad (5.24)$$

Here ρ_{xx} is simply the total resistance plus the $\Delta\rho$ AMR discussed theoretically in the previous section. The other component of resistivity, ρ_{yx} , is known as the **planar Hall effect (PHE)**, which is an off-diagonal ‘‘Hall’’ component of the experimentally measured in-plane magnetoresistance resulting from the anisotropic resistance [92, 93]. This component is *symmetric* with respect to the magnetic field and has a 180° periodicity. On the other hand, the ordinary Hall contribution resulting from a misalignment and out-of-plane field B_{\perp} , is antisymmetric with respect to B and can be easily removed from the PHE signal by utilising the data in both positive and negative B . As such, measurement of the PHE gives the genuine amplitude of $\Delta\rho = \rho_{\parallel} - \rho_{\perp}$ with no spurious contribution from misalignment.

5.4.2 Surface characterisation using the PHE

The PHE and AMR as a function of angle are shown in Fig. 5.5. We see from both the PHE and AMR as a function of field, Fig. 5.5(a-b), that they are consistent with our theory. In particular the anisotropy of resistivity is such that $\rho_{\parallel} > \rho_{\perp}$, so that a parallel current and field has a higher resistivity than in the perpendicular direction. The amplitude of the PHE is also shown in Fig. 5.5 and should be compared to the same amplitude that was theoretical calculated in Fig. 5.3. The experimentally measured PHE amplitude shows a clear two-peak structure, as predicted by our theory. There is a large peak for negative gate voltages and a smaller peak for positive gate voltage, suggesting that the impurities do not respect particle-hole symmetry.

We can see that the experimental results from Fig. 5.5 are in excellent qualitative agreement with our theory in Fig. 5.3. It is also, however, possible to use our theory to obtain quantitative information about the characteristics of impurities on our $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ surface. In particular, we have seen that the peaks in the anisotropic resistivity arise from peaks in the spin-flip scattering rate $\text{Im}\Sigma_{12}$, with a location set by the width of the impurity dominated regime $\mu \sim \pm\Gamma_0$. Hence, the location of μ associated with a peak gives a measure of Γ_0 . To estimate the chemical potential μ we can use the Hall data – which

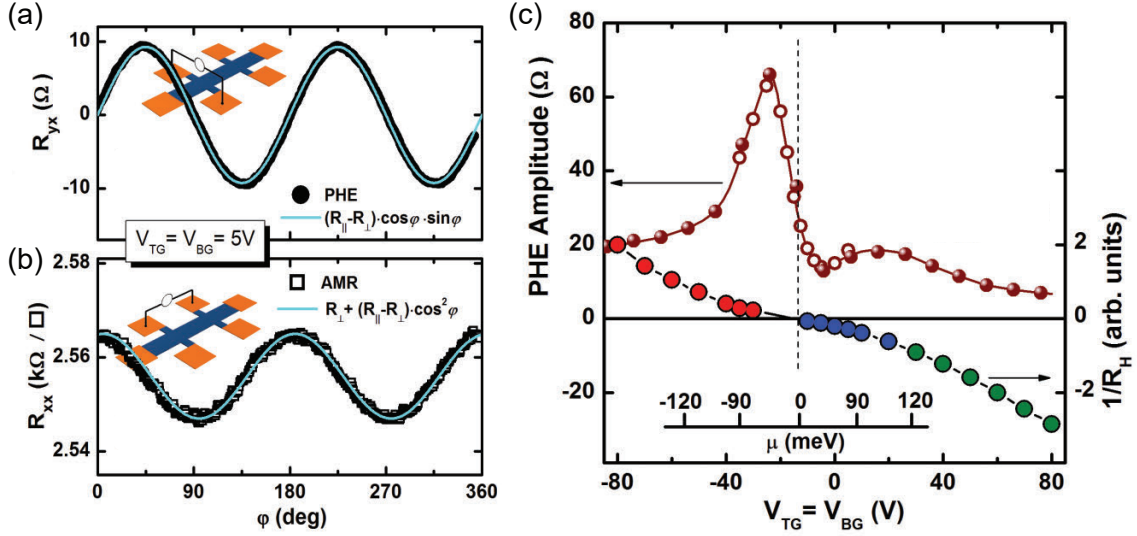


Fig. 5.5 **AMR and PHE in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$** : (a) Planar Hall effect as a function of angle showing the expected $\cos \varphi \sin \varphi$ due to anisotropy of magnetoresistance $\rho_{\parallel} > \rho_{\perp}$. (b) The magnetoresistance as a function of angle. Again a clear anisotropy can be seen with $\rho_{\parallel} > \rho_{\perp}$ due to the increased spin-flip scattering. The amplitude of the AMR is, however, susceptible to misalignment effects. (c) (left axis) the Planar Hall effect amplitude as a function of gate voltage, showing a clear two peak structure expected from theory. (right axis) The Hall conductance $1/R_H$ as a function of gate voltage. This is proportional to the charge density and so can be used to calculate the resulting change in chemical potential, as shown on the bottom axis.

directly measures the charge density which we have seen $n(\mu) \sim \mu^2$. Estimating the Fermi velocity as $3.9 \times 10^5 \text{ m/s}$ [54, 79] (see Fig. 5.5c), we obtain for our sample $\Gamma_0 \sim 50 \text{ meV}$, corresponding to a transport mean free path v_F/Γ_0 of $\sim 70 \text{ \AA}$.

Furthermore, we saw that the amplitude of the resistivity anisotropy is set by the square of the ratio of the spin-flip and non-spin-flip scattering rate, as in Eq. (5.20). From the $\sim 1\%$ size of the anisotropy at 14 T we can determine that the spin-flip scattering rate $\text{Im}\Sigma_{12}$ is $\sim 10\%$ of all scatterings, in other words the spin-flip transport mean free path is $\sim 700 \text{ \AA}$. We therefore see that the PHE cannot only give good qualitative agreement with theory, but a combination of the two allows us to obtain realistic values for the transport characteristics of the surface.

Summary and outlook

In this chapter we saw that it was possible to lift a topological insulator's protection from back-scattering, in a manner which retained the Dirac physics of the surface states, by applying an in-plane magnetic field. Whilst for the surface states the magnetic field only led to a shift in momentum-space of the position of the Dirac point, we argued that in the vicinity of impurities random magnetic fields parallel or anti-parallel to the magnetic

5.4. Anisotropic MR and planar Hall effect in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$

field would develop. These random fields enable spin-flip scattering when the spin has a component perpendicular to the field which leads to an anisotropy of resistance, with the resistance parallel to the field larger than the resistance perpendicular to it. By calculating the conductivity using linear response theory, including vertex corrections, we found that the amplitude of this resistivity anisotropy had a characteristic two peak structure as a function of chemical potential. Experimental measurements of the magnetoresistance in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ find an anisotropy of resistivity, which we argued was best measured using the planar Hall effect because it enables any sample misalignment effects to be eliminated. The amplitude of the planar Hall effect in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ also showed a two peak structure and by comparing theory and experiment we were able to ascertain the mean free path and spin-flip mean free path of the electrons on our $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ surface.

The controllable lifting of a topological insulator's protection presents many opportunities to further probe this effect. The transport experiments in this chapter were a direct measurement of the resistivity anisotropy but not of the lifting of protection. We have already seen in chapter 3 that it is possible to directly measure the protection from backscattering using quasi-particle interference. Unfortunately the magnetised impurities from this chapter will behave like the magnetic impurities discussed in chapter 3: Whilst there will be a response of the LDOS it will be opposite for opposite spins, therefore leading to no overall change in the LDOS. This is not true, however, if one were to measure the spin-resolved LDOS with an in-plane magnetic field applied, where we would expect a peak to grow with increasing magnetic field strength is increased and the spin-flip scattering rate increases.

It is also interesting to consider the effect such magnetised impurities will have on topological insulator nanowires (see next chapter). This is of particular interest because it is necessary to thread half a flux quantum along such a wire in order to close its gap, meaning that some of the most interesting physics in nanowires *requires* a magnetic field.

Quantum confinement in topological insulator nanowires

The properties of a topological insulator's surface state change dramatically when confined to a finite geometry. For instance consider a cylindrical wire of TI, due to the π Berry phase, the surface wave function has to obey **anti**-periodic boundary conditions around such a wire. This boundary condition results in a half integer quantisation of total angular momentum and a discrete, spin-degenerate, sub-band structure in terms of the momentum k along the wire. In this chapter we discuss the consequences this quasi-1D geometry has on transport properties along the TI wire in the diffusive limit, that is, when the mean free path is shorter than the length of the wire. In particular we will focus on these properties as a function of chemical potential, which we have seen in previous chapters can be easily changed experimentally by applying a gate voltage. We will then compare our theoretical predictions with such an experiment on very thin wires of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$.

The layout of this chapter is as follows: First we will discuss the consequences of a finite geometry on the band-structure of the TI surface and calculate the density of states and charge density. Having established this we will proceed to add disorder and consider the conductivity as a function of chemical potential. In the last section of this chapter we discuss how such wires can be created experimentally and show that our theoretical predictions agree well with experiment.

This chapter is based on the preprint “Quantum-confined surface channels in bulk-insulating topological-insulator nanowires” and its corresponding supplementary material.

6.1 Surface states on a wire

6.1.1 A topological insulator wrapped up

Recall from chapter 1 that the most general form of the Hamiltonian for the bulk of a 3D topological insulator is a 4×4 matrix denoting the 4 states of the p_z orbitals on, for instance, a Bi atom and other element (e.g Se) is

$$H = \begin{pmatrix} M & Ap_z & 0 & Ap_- \\ Ap_z & -M & Ap_- & 0 \\ 0 & Ap_+ & M & -Ap_z \\ Ap_+ & 0 & -Ap_z & -M \end{pmatrix} = \tau_z \mathbb{1} M + \tau_x \boldsymbol{\sigma} \cdot \mathbf{p}, \quad (6.1)$$

Chapter 6. Quantum confinement in topological insulator nanowires

where $k_{\pm} = p_x \pm ip_y$. We showed in chapter 1 that for the mass term to vanish on the surface of the half-volume $z < 0$, the 4-component spinor $\{\phi, \xi\}^T$ had to satisfy $\phi = i\sigma_z \xi$. However there was nothing special in our choice of $\hat{\mathbf{z}}$ as the normal direction and it is easy to see by symmetry [94, 95] that the spinors on a surface with a local normal $\hat{\mathbf{n}}$ in any direction must simply satisfy the condition $\phi = i(\mathbf{n} \cdot \boldsymbol{\sigma})\xi$. This leads to the general surface Hamiltonian for a curved surface

$$H_{\text{surf}} = v_F \mathbf{n} \cdot (\boldsymbol{\sigma} \times \mathbf{p}). \quad (6.2)$$

In particular we are interested in the case of a cylindrical surface which is infinite in the x -direction and of circumference $2\pi R_w$ in the $y - z$ plane. Such a surface has normal $\mathbf{n} = \{0, \sin(\phi), \cos(\phi)\}$ and the Hamiltonian depends only on momentum around the wire, $\hat{p}_{\phi} = \hat{L}_{\phi}/R_w = -i\hbar\partial_{\phi}/R_w$, and momentum along the wire, $\hat{p}_x = -i\hbar\partial_x$,

$$H_{\text{surf}} = i\hbar v_F \left(\frac{\sigma_x}{R_w} \partial_{\phi} - (\sin(\phi)\sigma_z - \cos(\phi)\sigma_y) \partial_x \right). \quad (6.3)$$

This can be simplified by applying a spinor-rotation by ϕ about the x -axis, *i.e.*, the unitary transformation $U_x(\phi) = \exp(-i\phi\sigma_x/2)$. The result is the Hamiltonian

$$H_0 = i\hbar v_F \left(\frac{\sigma_x}{R_w} \partial_{\phi} - \sigma_y \partial_x \right). \quad (6.4)$$

Importantly, since spinor rotations are 4π periodic, the eigenfunctions of this Hamiltonian must obey anti-periodic boundary conditions, resulting from $U_x(\phi) = -U_x(\phi + 2\pi)$. The wave functions can therefore be written in the general form $\psi_{\ell}(k) = e^{i(kx + \ell\phi)}\xi$, where the anti-periodic boundary conditions require the total angular momentum to satisfy $\ell = \pm 1/2, \pm 3/2, \dots$ and ξ is a spinor encoding the spin-structure of the surface state within this rotated system.

Since ℓ is quantised as a half integer, due to the finite geometry along the circumference of the wire, the eigenenergies of the Hamiltonian are also quantised and given by

$$E_{\ell, \pm}(k) = \pm \hbar v_F \sqrt{k^2 + \left(\frac{\ell}{R_w} \right)^2}. \quad (6.5)$$

The result is a series of doubly degenerate bands with a gap at zero chemical potential μ , as is shown in Fig. 6.1, the band edge of each band given by $\varepsilon_{\ell, \pm} = \pm \hbar v_F \ell / R_w$. For simplicity in what follows $\varepsilon_{\ell} = \varepsilon_{\ell, +}$.

In the presence of a flux Φ threaded along the wire, ∂_{ϕ} is replaced by $\partial_{\phi} - i\eta$ with $\eta = \Phi/\Phi_0$ [95]. Although not of direct relevance here, the magnetic field is able to restore a non-degenerate 1D gapless mode for half integer η [96, 97]. This makes topological insulator nanowires an excellent candidate to host Majorana zero-modes when coupled to a superconductor [98–100].

6.1. Surface states on a wire

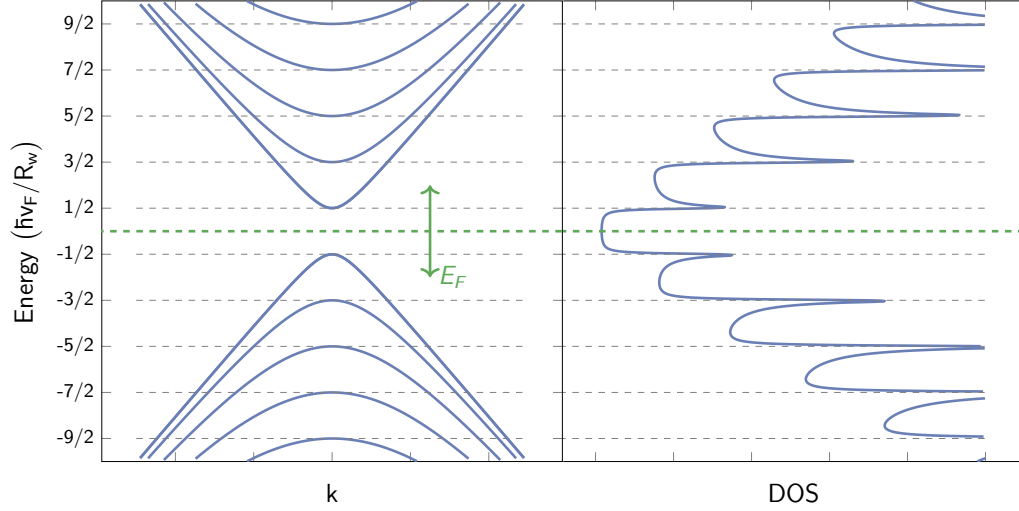


Fig. 6.1 **The sub-bands of a topological insulator wire:** (Left) As a function of k along the wire. Each band is doubly degenerate and the band-edge occurs at $k = 0$ and $E = \varepsilon_{\ell, \pm} = \pm \hbar v_F \ell / R_w$, where ℓ is total angular momentum around the wire. (Right) The density of states tracks that which is expected from a 2D Dirac Hamiltonian and divergences at the edge of each sub-band, ε_{ℓ} , in the picture these divergences are smoothed out by the presence of a finite amount of disorder.

6.1.2 Density of states and charge density

We now calculate the density of states and charge density of the nanowire without disorder, both quantities will be important for our theoretical understanding of the transport properties of the wire and comparison to experiment.

The free retarded Green's function – which, as we have seen previously, is a matrix – has the form

$$\mathbf{G}^0(\mu, k, \ell) = (\mu \mathbb{1} - H_0(k, \ell) + i\delta \mathbb{1})^{-1}, \quad (6.6)$$

where we take the matrix inverse. Here δ acts as some overall background scattering rate resulting from other degrees of freedom (e.g. impurity bands) not described by our Hamiltonian. It has the effect of broadening the divergence of the density of states at the bottom of each sub-band.

The local Green's function of each ℓ resolved sub-band can be calculated analytically and is given by

$$\Delta(\mu, \ell) = \frac{1}{2\pi R_w v_F} \int \frac{dk}{2\pi \hbar} \mathbf{G}^0(\mu, k, \ell) = -\frac{(\mu + i\delta) \mathbb{1} - \varepsilon_{\ell} \sigma_x}{4\pi v_F R_w \hbar \sqrt{\varepsilon_{\ell}^2 - (i\delta + \mu)^2}}. \quad (6.7)$$

From this local Green's function we can obtain the local density of states

$$\rho(\mu) = -\frac{1}{\pi} \sum_{\ell=\pm\frac{1}{2}, \pm\frac{3}{2}, \dots} \text{Im} \{ \text{Tr} \Delta(\mu, \ell) \} = \frac{1}{2\pi^2 R_w v_F \hbar} \sum_{\ell=\pm\frac{1}{2}, \pm\frac{3}{2}, \dots} \text{Im} \left\{ \frac{\mu + i\delta}{\sqrt{\varepsilon_{\ell}^2 - (i\delta + \mu)^2}} \right\}, \quad (6.8)$$

which, as should be expected, tracks the 2d Dirac density of states $\rho^{2d}(\mu) = |\mu|/2\pi v_F^2 \hbar^2$. In stark contrast to an infinite surface, however, the density of states a TI nanowire diverges when a band-edge is reached, $\mu \rightarrow \varepsilon_\ell$, and a new sub-band is added to the system. The density of states is shown in Fig. 6.1,

An important quantity for our comparison to experiment will be the charge density along the wire. This is because, as discussed in chapter 4, it is the charge density which can be altered in proportion to gate voltage. It is simple to calculate the charge density by integrating the density of states over energy, such that

$$n(\mu) = 2\pi R_w \int_0^\mu d\mu' \rho(\mu') = -\frac{1}{\pi v_F \hbar} \text{Im} \left\{ \sum_{\ell=\pm\frac{1}{2}, \pm\frac{3}{2}, \dots} \sqrt{\varepsilon_\ell^2 - (\mu + i\delta)^2} \right\}. \quad (6.9)$$

We see that $n(\mu)$ is zero at the Dirac point. For $\mu > 0$ it counts the contribution of all bands above the Dirac point up to the chemical μ . Correspondingly, for $\mu < 0$ it counts the charge of all holes contributed by bands below the Dirac point, in other words $n(\mu)$ is negative for $\mu < 0$.

6.2 A disordered TI nanowire

We add disorder to the Hamiltonian Eq. (6.4) by introducing point-like impurity potentials located at random positions \mathbf{r}_i , $V_i(\mathbf{r}) = u_0 \mathbb{1} \delta(\mathbf{r} - \mathbf{r}_i)$, where \mathbf{r} and \mathbf{r}_i are 2d coordinates on the surface of the wire. Since the impurities are spin-diagonal, the above unitary transformation $U(\phi)$ to convert H_{surf} to the simpler H_0 has no effect on the disorder potential and we can write the disordered system in terms of the simplified Hamiltonian in the rotated spin-basis

$$H = H_0 + \sum_i V_i(\mathbf{r}) \mathbb{1}. \quad (6.10)$$

6.2.1 T-matrix approximation

The full disorder averaged Green's function for each sub-band defined by ℓ is

$$\mathcal{G}(\mu, k, \ell) = (\mu \mathbb{1} - H_0(k, \ell) + \Sigma(\mu))^{-1}, \quad (6.11)$$

where $\Sigma(\mu)$ is the full self-energy from all scattering processes, including disorder and the other scattering processes modelled by the factor $i\delta$ (see above). After disorder averaging, the self-energy $\Sigma(\mu)$ is independent of both ℓ and k due to the local nature of the impurities.

To approximate the self-energy contribution due to disorder scattering we use the full T-matrix calculated at first order in impurity density n_{imp} (i.e. the non-crossing approximation). As seen in the previous chapters, within this approximation the self-

energy, which is also a matrix, is

$$\begin{aligned}\Sigma(\mu) &= n_{\text{imp}} u_0 \left\langle \left(\mathbb{1} - \frac{u_0}{2\pi R_w} \sum_{\ell=\pm\frac{1}{2}, \pm\frac{3}{2}, \dots} \int \frac{dk}{2\pi\hbar v_F} \mathbf{G}^0(k, \mu, \ell) \right)^{-1} \right\rangle_{\text{imp}} \\ &= n_{\text{imp}} u_0 \left(\mathbb{1} - u_0 \sum_{\ell=\pm\frac{1}{2}, \pm\frac{3}{2}, \dots} \Delta(\mu, \ell) \right)^{-1}.\end{aligned}\quad (6.12)$$

For small δ , the local Green's function $\Delta(\mu, \ell)$ from Eq. (6.7), only has a substantial imaginary part for $|\mu| > \varepsilon_\ell$, this means that a conductivity channel only contributes to the self-energy when the corresponding sub-band has an occupation of electrons (holes) for the upper (lower) Dirac cone. Further the off-diagonal components of Σ are exactly zero for diagonal impurities since the off-diagonal contributions of positive and negative ℓ cancel.

The self-energy can also be calculated self-consistently by replacing $\mu \rightarrow \mu + \Sigma_{11}$ and $\varepsilon_\ell \rightarrow \varepsilon_\ell + \Sigma_{12}$, where Σ_{11} and Σ_{12} are the diagonal and off-diagonal components, respectively. Self-consistency has a similar impact on the self-energy as δ , that is, the breadth of the peaks in scattering rate become broadened in proportion to the self-energy itself. Ultimately this means that peaks at high sub-band index become “washed out” when the self-energy is of the same order of magnitude as the sub-band spacing. This is not of relevance for the experimental situation discussed here where the mean free path is several times the wire radius, but would limit the number of peaks seen in thicker or dirtier wires where the ratio of mean free path and radius can be substantially smaller.

6.2.2 Conductivity

As we have seen in the last 2 chapters, the conductivity is found using the Kubo formula to calculate the current-current correlation function in linear response. In this instance we neglect vertex corrections since, as in the last chapter, they are likely only have a quantitative effect on overall conductivity. Within this scheme the DC conductivity is given by

$$\sigma = \frac{e^2 v_F^2 \hbar^2}{\pi} \frac{1}{2\pi R_w} \sum_{\ell=\pm\frac{1}{2}, \pm\frac{3}{2}, \dots} \int d\omega n'_F(\omega) \int \frac{dk}{2\pi\hbar} \text{Tr} \left(\sigma_y \mathbf{G}(\omega, k, \ell) \sigma_y \mathbf{G}^\dagger(\omega, k, \ell) \right). \quad (6.13)$$

We can also now consider the contribution of each sub-band ℓ to the conductivity by writing total conductivity as a sum over each contribution

$$\sigma = \sum_{\ell=\pm\frac{1}{2}, \pm\frac{3}{2}, \dots} \sigma_\ell(\mu). \quad (6.14)$$

Due to the simple 1d-like nature of the bands the contribution σ_ℓ can in fact be calculated entirely analytically

$$\begin{aligned}\sigma_\ell(\mu) &= \frac{e^2 v_F}{(2\pi)^2 R_w} \frac{1}{\tilde{\mu} \Gamma_{11} - \tilde{\varepsilon}_\ell \Gamma_{12}} \text{Im} \left\{ \frac{\lambda}{\sqrt{\xi}} - \frac{\lambda^\dagger}{\sqrt{\xi^\dagger}} \right\} \\ &\approx \theta(\mu^2 - \varepsilon_\ell^2) \frac{e^2 v_F}{2\pi^2 R_w |\Gamma_{11}|} \frac{\mu^2 - \varepsilon_\ell^2}{\mu \sqrt{\mu^2 - \varepsilon_\ell^2}},\end{aligned}\tag{6.15}$$

where $\mathbf{\Gamma} = \text{Im}\mathbf{\Sigma}$ is the (spin-resolved) scattering rate, $\tilde{\mu} = \mu + \text{Re}\Sigma_{11}$, $\tilde{\varepsilon}_\ell = \varepsilon_\ell + \text{Re}\Sigma_{12}$, and the factors $\xi = -(\mu + i\Gamma_{11})^2 + (\varepsilon_\ell + i\Gamma_{12})^2$ and $\lambda = -\tilde{\varepsilon}_\ell^2 - i\tilde{\varepsilon}_\ell \Gamma_{12} + i\tilde{\mu} \Gamma_{11} + \tilde{\mu}^2$ ensure that a channel only contributes to conductivity when $|\mu| \gtrsim \varepsilon_\ell$. The approximation in (6.15) is valid for small self-energies and $\theta(\mu^2 - \varepsilon_\ell^2)$ is the Heaviside function.

When chemical potential is varied such that a new sub-band is added, we see from Eq. (6.15) that there exist two competing effects: 1) The new conduction channels parallel to the wire provide an additional contribution to the conductivity which can *decrease* the resistivity as more charge can be transported and 2) The addition of extra back-scattering channels reduce the conductivity of all sub-bands, including those which are already occupied causing an *increase* the resistivity.

To understand the overall outcome of these effects, which is shown schematically in Fig. 6.2, one should first consider when the chemical potential is tuned to the sub-band edge, ε_ℓ . Here the new conduction channel has a very small occupation and thus a negligible contribution to the conductivity. In contrast all sub-bands that are already occupied experience a sizeable reduction in their conductivity since Γ is substantially larger and they are able to scatter to the newly available backward moving channels. As such, the overall effect of the addition of a new band is a decrease in the total conductivity close to ε_ℓ or, equivalently, a peak in the total resistivity. As chemical potential is varied away from ε_ℓ , deeper into the sub-band, the total conductivity is able to benefit from the new conductance channel, allowing the conductivity to increase away from its minimal value close to ε_ℓ . This increase continues until the next band edge $\varepsilon_{\ell+1}$ is reached, at which point overall conductivity once again decreases due to increased scattering. The result is a series of equally spaced dips in total conductivity (peaks in resistivity) as a function of chemical potential with each conductivity dip centered at the position of the sub-band edge ε_ℓ .

6.2.3 Dilution of impurity strength

An important parameter for our theory is u_0 , the strength of the impurity potential. In Fig. 6.3 we show the resistivity as function of chemical potential for weak, intermediate and strong impurity potentials. Only for weak impurity potentials, $\frac{u_0}{\hbar v_F R_w} \ll 1$, peaks occur at the band edges and particle- and hole doping are almost equivalent. For intermediate impurity strength, $\frac{u_0}{\hbar v_F R_w} \sim 1$, the curve is highly asymmetric. For $\frac{u_0}{\hbar v_F R_w} \gg 1$, the particle-hole symmetry is restored but instead of peaks one obtains dips.

6.2. A disordered TI nanowire

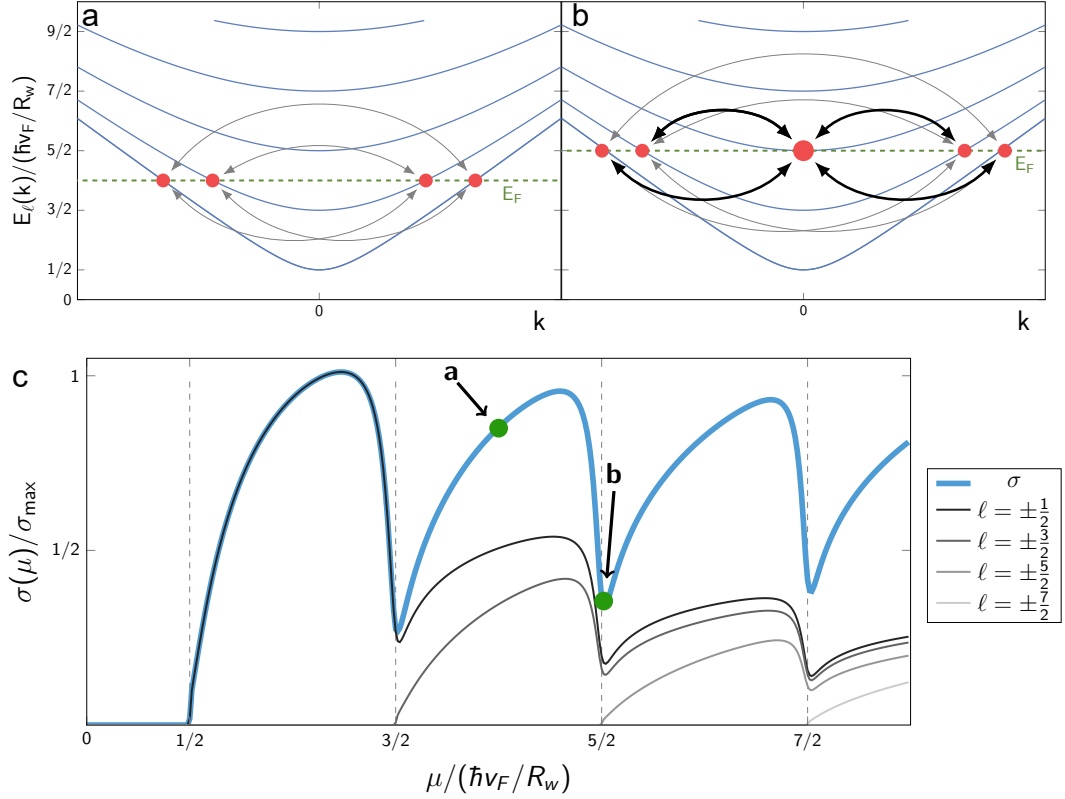


Fig. 6.2 **Conductivity of a disorder TI nanowire:** **a, b**, Scattering channels (thin gray arrows) for two positions of μ . When μ is at the bottom of a sub-band, as in panel **b**, all other sub-bands scatter at an enhanced rate into the new sub-band due to its diverging density of states, leading to a pronounced minimum in the conductivity. **c**, Theoretically-calculated conductivity as function of μ . Thin black lines display the contribution of each sub-band labeled by $\ell = \pm\frac{1}{2}, \pm\frac{3}{2}, \pm\frac{5}{2}, \pm\frac{7}{2}$, which add up to give the total conductivity (thick blue line). The conductivity of all channels shows pronounced minima at $\mu = \ell\hbar v_F/R_w$, when the chemical potential touches the bottom of a new sub-band.

We will see that only weak u_0 fit the data from experiment. This is expected because the wave-function of the conduction channel wraps around the circumference of the wire, effectively diluting the impurity by a factor $1/(2\pi R_w)$. To be precise: Weak scattering can be defined by $u_0\bar{\rho} \ll 1$, where $\bar{\rho} \sim \frac{1}{2\pi v_F \hbar R_w}$ is the typical density of states in the system for small ℓ . Assuming that the width of the scattering potential is of the order of the lattice constant a , the weak scattering limit is reached when the amplitude, $V_0 = u_0/a^2$, of the scattering potential fulfils $V_0 \ll \frac{v_F}{a\hbar} \frac{2\pi R_w}{a}$. Since in any experiment the circumference of a TI wire is many lattice constants – e.g. in the experiments discussed later $\frac{R_w}{a} \sim 100$ – the weak scattering limit is generically realised.

6.2.4 Matrix elements between sub-bands

While the main contribution to the peaks at the sub-band edges arises from the diverging density of states, this effect is further enhanced by a matrix-element effect closely related to

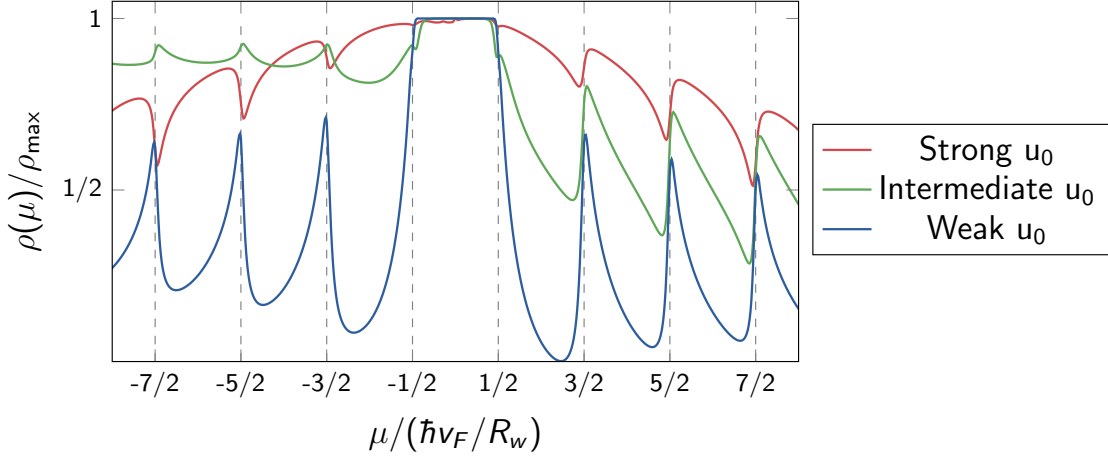


Fig. 6.3 **Resistivity as a function of chemical potential in the weak, intermediate, and strong scattering regimes:** A weak scatterer, $\frac{u_0}{\hbar v_F R_w} = 0.1$, $n_{imp} = 20/R_w^2$, is shown as the blue curve. The green curve, $\frac{u_0}{\hbar v_F R_w} = 1$, $n_{imp} = 2/R_w^2$, describes scatterers of intermediate strength, the red curve, a strong scatterer $\frac{u_0}{\hbar v_F R_w} = 20$, $n_{imp} = 0.1/R_w^2$. Strong scattering is capable of turning resistivity peaks at sub-band edges into resistivity dips, however such a situation is unlikely (see text).

the topological protection of TI surface states that has been discussed in previous chapters. Here the spin orientation is locked to the propagation direction and scattering from \mathbf{k} to $-\mathbf{k}$ is prohibited by time-reversal symmetry. As we saw above, in nanowires, $k_y = \ell/R_w$ is quantized, and the 1D bands are doubly degenerate as for each $k_x = k$, two transverse momenta, $\pm k_y$, are possible. For $k_x \gg k_y$, however, the 2D vectors $\mathbf{k} = (k_x, \pm k_y)$ and $\mathbf{k} = (-k_x, \pm k_y)$ are almost antiparallel, leading to a suppression of the scattering rate by the factor $(k_y/k_x)^2$. This matrix element effect (encoded in the 2×2 matrix structure of the T-matrix) substantially suppresses backscattering among occupied channels with $\ell/R_w \ll k_x$ relative to the scattering to a newly opened channel with a large k_y .

6.3 Quantum confinement in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ nanowires

We now discuss the experimental realisation of topological insulator nanowires with the intention of comparing our theoretical transport predictions to those found in experiments on $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$. In order to achieve quantum confinement of the surface states which results in sizeable conductivity effects, any experimental nanowires should be: 1) Bulk insulating so that the transport effects of the surface are dominant and not “washed out” by the conductivity of the bulk and 2) The wires must be narrow so that the sub-bands are well defined in energy, in practice this means that the energy gap between sub-bands should be much larger than the self-energy of the Dirac quasiparticles, which we will see for $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ requires wires thinner than ~ 100 nm.

Largely due to these restrictions, previous measurements of TI nanowires have only been able to indirectly probe the quantized Dirac sub-bands far from the Dirac point [101–106], mostly from oscillations of resistivity as a function of the magnetic flux Φ threaded through

6.3. Quantum confinement in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ nanowires

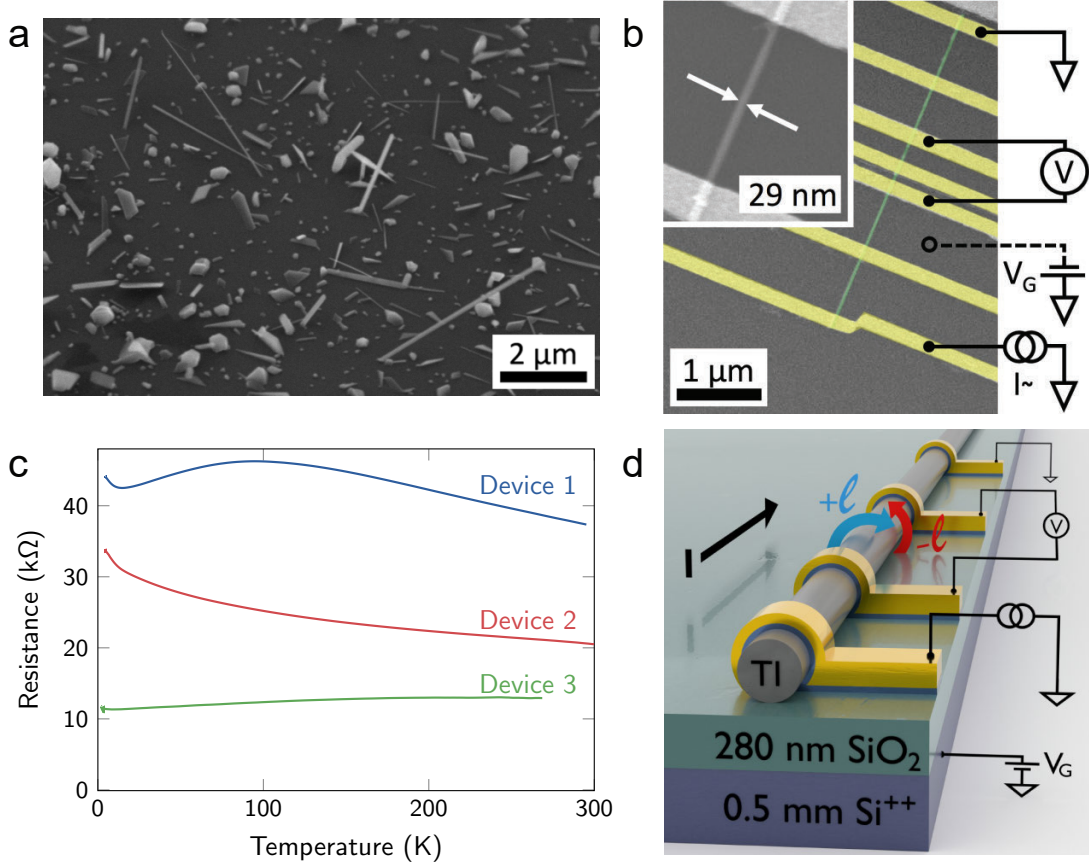


Fig. 6.4 **Experimental set-up:** a) Candidate nanowires from VLS growth b) An SEM image of device set-up showing the contacts and gate attached to the wire c) Resistance as a function of temperature. Although for some devices e.g. device 3 the resistance looks potentially metallic, all 3 devices are bulk-insulating as shown by their gate voltage dependence. d) An overview of the full experimental set-up.

the wire. This effect arises because, as discussed above, the sub-bands can be shifted in energy by a magnetic field along the wire and so there is a periodic change in the number of occupied sub-bands at a set chemical potential [95]. Such oscillations have a period of one flux quantum $\phi_0 = h/e$ since for integer ϕ/ϕ_0 the doubly degenerate sub-band structure is restored.

6.3.1 Bulk insulating nanowires

Vapour-liquid-solid growth of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ using Au nanoparticles can obtain nanowires with a high crystalline quality and constant diameter between 20 to 100 nm and a length of several μm . Some of the candidate wires are shown in Fig. 6.4a. After growth these wires are placed on a gate and contacts are attached to create devices which enable four terminal measurement of resistivity. In total we will discuss 3 such devices each with radii R_w of 21.5, 20, and 14.5 nm for devices 1, 2, and 3, respectively. Device 2 is shown in the

scanning electron microscope image in Fig. 6.4b, which also shows its very thin width of 29 nm. A schematic giving a full overview of the experimental set-up is shown in Fig. 6.4d

In Fig. 6.4c resistance R as a function of temperature T is shown for each of the 3 devices. These present what looks like insulating or metallic behaviour depending on the device. In fact, all three samples are bulk-insulating and have a chemical potential in the vicinity of the Dirac point; this is indicated by the clear resistance maximum in their gate-voltage dependences (see Fig. 6.5), showing that the Dirac point is crossed by varying gate voltage. The difference in the $R(T)$ dependence is likely explained by a slightly different electron density n of the three samples in the absence of gating. We find that the zero gate dopings are $n \approx -0.2, +0.4, +0.5 \text{ nm}^{-1}$ relative to the Dirac point for devices 1, 2, and 3.

6.3.2 Resistance of real nanowires

We now turn to the experimental feature of relevance for our theoretical discussion, namely the resistivity (or resistance) as a function of chemical potential or, in practice, gate voltage. The gate voltage dependence of the resistivity is shown for each of the 3 devices in Fig. 6.5. We start by discussing the experimental results and then what is required to compare our theory with these experimental observations.

The resistance traces of Fig. 6.5 show two distinct fluctuation features: Most prominent are semi-oscillatory features in the gate voltage with amplitudes $A_I \approx 5 \text{ k}\Omega$ (we call these type I). We will show these are the peaks in resistivity predicted by our theory due to sub-band crossings. Secondly, smaller fluctuations have amplitudes $A_{II} \approx 0.5 \text{ k}\Omega$ (type II) and are time dependent. Our theory does not provide an explanation for these type II oscillations. It is possible that they are related to the electron-hole puddles that arise due to the compensation of the wire [84, 85, 78].

On the theoretical side, as in previous chapter 4, it is important to understand exactly how gate voltage varies the chemical potential. We saw that the relation between gate voltage and chemical potential has two terms dependent on the Coulomb energy and chemical potential directly and can be expressed as $eV_G = \frac{e^2}{C_G}n(\mu) - \mu$, where C_G is the geometric capacitance of the device, $e < 0$ is the electron charge, μ the chemical potential, and $n(\mu)$ the electron density as calculated in Eq. (6.9). Once again, the Coulomb contribution dominates for the range of chemical potentials considered in the experiment. To be more precise, our analysis will show that a change of the gate voltage by 30 V leads to changes of the chemical potential by less than 100 meV. This means that the effect of the geometric capacitance dominates by more than two orders of magnitude and it is safe to approximate $V_G \approx \frac{1}{C_G}en(\mu)$. As such only one fit parameter, the geometric capacitance C_G , is required to set the positions of the peaks as a function of gate voltage¹.

¹In principle there is also a loose dependence of peak position on the impurity strength u_0 . In practice, however, as was argued previously, u_0 is weak meaning that particle-hole symmetry is approximately conserved or, in other words, the 1st Born approximation - which goes as $\sim u_0^2$ is independent of the sign of u_0 and thus particle-hole symmetric - is a good approximation to the full T-matrix

6.3. Quantum confinement in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ nanowires

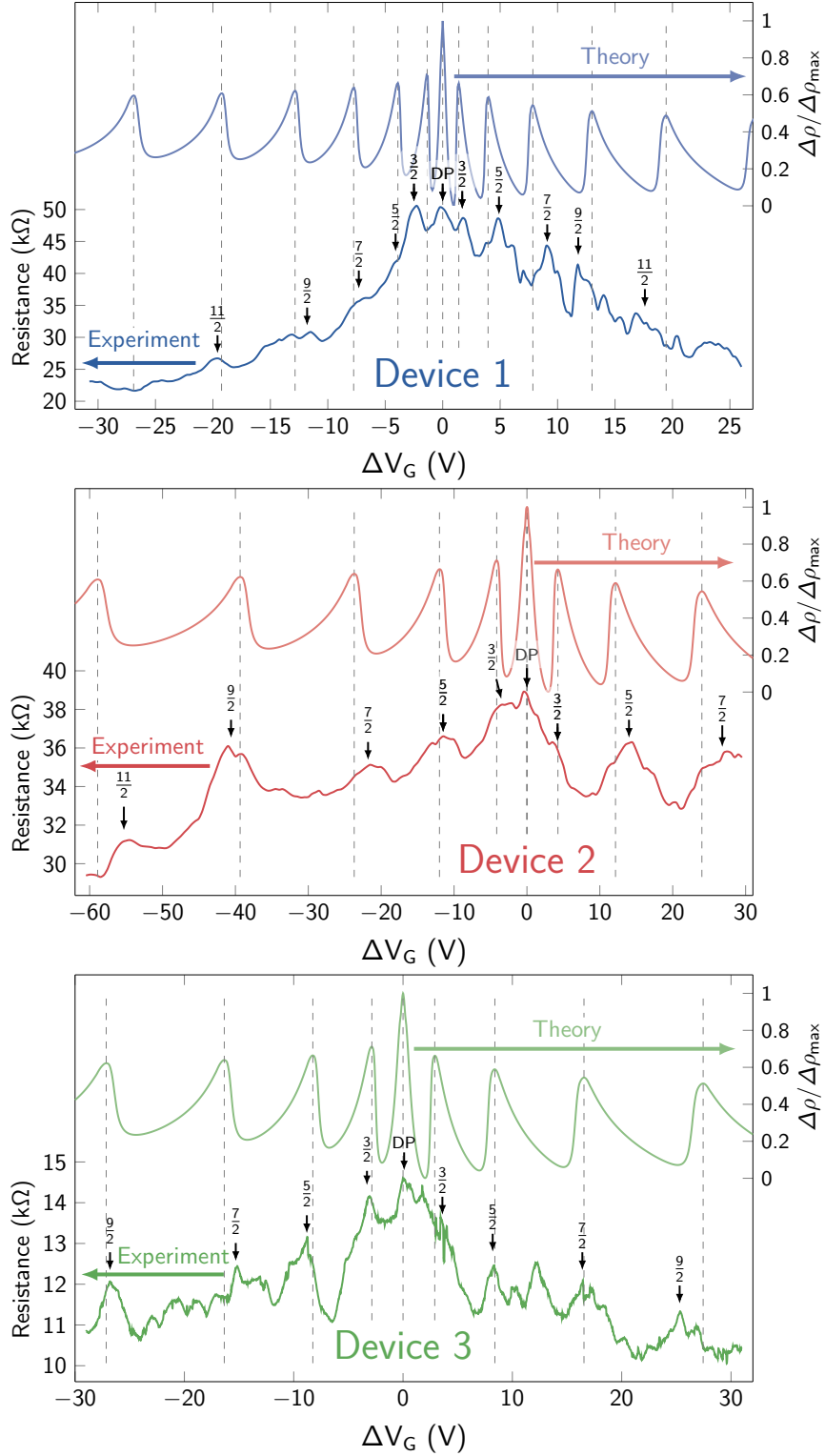


Fig. 6.5 **Gate voltage dependence of resistance for 3 different nanowire devices:** Lower curves show the V_G dependence of R observed in devices 1–3 at 2 K; $\Delta V_G = 0$ corresponds to the Dirac point of the TI surface state, which was achieved with V_G of -6.0 , 30.5 , and 19 V in devices 1, 2, and 3, respectively. Upper curves in panels are the theoretically-calculated resistivity. Pronounced maxima should arise at sub-band crossings (dashed lines).

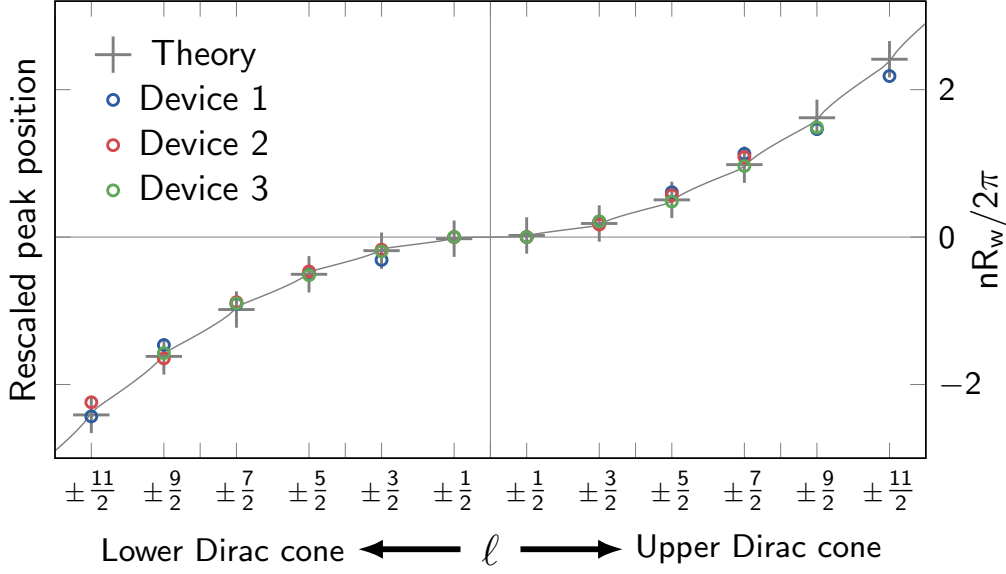


Fig. 6.6 **Theory vs experimental peak position:** Rescaled position of the resistivity maxima of the three devices ($\Delta V_G/V_0$, see main text) as function of the quantum number ℓ compared to the theoretically calculated n (in units of $2\pi/R_w$) at sub-band crossings.

It is, however, also possible to reproduce the *magnitude* of the resistivity effect. It turns out that a weak scatterer $u_0 = 0.1v_F\hbar R_w$, relatively dilute impurities $n_{\text{imp}} = 20/R_w^2$, and small Dirac quasi-particle broadening $\delta = 0.025v_F\hbar/R_w$ achieve the ~ 5 nm/k Ω of the fluctuations in wire conductivity found experimentally. These values are used in Fig. 6.5. To give a feeling of the scale of these parameters after the first couple of peaks they correspond to a typical mean free path of the order of $l \sim 100$ nm, which is smaller than, or approximately equal to, the wire length for all ℓ .

We now discuss the main result of this section, the fit of experiment and theoretical resistivity/resistance as a function of gate voltage as shown in Fig. 6.5. We find that the experiment fits our theory well for the capacitances per length of wire: $C_G = 5.8, 2.0$, and 4.1 pF/m for devices 1, 2, and 3, respectively. In this Fig. 6.5, the position of the peaks are labelled by the angular momentum quantum number ℓ of the added channel. When the chemical potential reaches the bottom of the first electron or the top of the first hole band ($\ell = \pm\frac{1}{2}$), the charge density is approximately zero in both cases and therefore there is only a single peak in the center for $\ell = \pm\frac{1}{2}$. For large ℓ , the peak position scales with ℓ^2 .

It is clear from Fig. 6.5 that our theory can reproduce well the most prominent features of the experiment, in particular the super-linear dependence of the locations of the type I peaks in the $R(V_G)$ traces. We can actually visualise this agreement further, since the experimental location of the peaks (indicated by arrows in Fig. 6.5) can be rescaled to a dimensionless measure of charge density by the scaling factor $V_0 = \frac{2\pi e}{R_w C_G}$. These rescaled positions, shown as function of the sub-band index ℓ in Fig. 6.6, have the advantage that the peak positions are entirely independent of device characteristics and so can be compared to the theoretically calculated electron density n at the peak position – which in Fig. 6.6 as shown as a grey line in units of $2\pi/R_w$. It can be seen that the super-linear behaviour

6.3. Quantum confinement in $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ nanowires

in the experimentally obtained V_G -dependent sub-band crossings is in excellent agreement with theory, therefore showing that the experiment was able to resolve a direct signature of the quantum-confined TI surface states.

Summary and Outlook

The observation of quantum confined surface states in the vicinity of a TI nanowire's Dirac point presents a wide range of experimental and theoretical opportunities. The most natural next step is observation of the coupling of the sub-bands to a magnetic field and closing of the gap at the Dirac point for half-integer flux quanta. The ultimate goal of this would be access to Majorana zero modes predicted to appear at the end of the wire when in proximity to a superconductor. If successful this would eventually open the tantalising possibility of Majorana based quantum computers built upon TI nanowires. Additionally, access to thin, bulk-insulating, and low electron density wires enables the experimental investigation of many mesoscopic phenomena such as universal conductance fluctuations and Klein tunnelling.

Since most previous theoretical studies [96, 107, 106] have considered semi-ballistic transport in TI nanowires and are largely numerical, there are also several outstanding questions on the theoretical side in regard to diffusive transport: What happens when a magnetic field is applied? What are the effects of magnetic (or magnetised impurities)? Can we understand the full “phase diagram” of magnetic field, impurity strength, and chemical potential? Moving away from our transport probe of the sub-bands, what are the effects of puddles and to what extent can these explain the smaller type II oscillations seen by experiment? Throughout our theory we assumed a circular wire but what is the impact, if any, of device geometry? Finally since experimental efforts to produce high quality, thin, bulk-insulating nanowires are still very much in their infancy it is highly likely that surprises hide around the corner, surprises that will require theoretical answers.

Disordered Dirac materials in a magnetic field

In this chapter we will consider the impact that a magnetic field has on Dirac fermions. In particular we will see that Dirac fermions are much more sensitive to magnetic fields than Schrödinger fermions. Unlike previous chapters we move away from topological insulators and focus on 3D Dirac materials, although several aspects – e.g. the increased sensitivity to a magnetic field – also apply to 2d Dirac materials such as topological insulators and graphene. Nonetheless, of particular interest in both this chapter and the next will be that 3d Dirac materials, as discussed in chapter 1, have a Hamiltonian that makes use of all three Pauli matrices. Unlike their 2D counterparts they always possess a direction which is unaffected by the application of a magnetic field and two mutually orthogonal directions, perpendicular to the magnetic field where the wavefunctions form Landau levels.

Since Dirac fermions are more sensitive to magnetic fields [108] it is much easier to reach the lowest Landau level – the quantum limit – at experimentally realisable fields. Entering the lowest Landau level results in several fascinating phenomena due to the quasi-1d nature of the fermions in the lowest Landau level. In this chapter we will provide an overview of both the response to disorder and a brief overview of transport properties of Dirac fermions in a magnetic field. In the following chapter we will use the concepts developed here to propose a new phenomena in materials with multiple Dirac points.

7.1 Landau levels in 3D Dirac materials

In this section we show the impact that the application of a magnetic field has on 3D Dirac fermions. Since we will later be interested in materials with multiple Dirac points – which is anyway enforced when time reversal and inversion symmetry are present – we will consider the simple model low energy Hamiltonian of a Dirac semi-metal

$$H = \tau_z v_F (\boldsymbol{\sigma} \cdot (\mathbf{k} \pm \mathbf{K}_0)), \quad (7.1)$$

with two Dirac nodes located at $\pm \mathbf{K}_0$ in the Brillouin zone and a particularly simple Pauli-basis in terms of two sets of matrices σ and τ where a matrix Kronecker product results in a 4×4 matrix Hamiltonian. As we saw in chapter 1, each Dirac point is created from two Weyl Hamiltonians with opposite chirality. If time-reversal or inversion symmetry is broken the Dirac point can split into two Weyl points resulting in a separation of each chirality close to $\pm \mathbf{K}_0$. Clearly it *is* the case that time reversal is broken in the presence of a magnetic field, however in practice this splitting is small – of the order of the Zeeman

energy – and we will neglect it in the present discussion. Finally we note that it is clear the band cannot continue forever so this Hamiltonian can only be true for low energies with respect to the Dirac point (or equivalently $|\mathbf{k}| \ll |\mathbf{K}_0|$) and there is no requirement (depending on symmetries) for the matrices or v_F to be the same at both Dirac points.

We now discuss the Hamiltonian Eq. (7.1) in the presence of a magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$, where \mathbf{A} is the vector potential. We of course have a gauge choice to make for the vector potential $\hat{\mathbf{A}}$; the simplest choice is the Landau gauge $\mathbf{A} = (By, 0, 0)$, ensuring that the magnetic field lies in the z-direction and we pick the x-direction such that the Dirac node separation $\mathbf{K}_0 = (k_{0x}, 0, k_{0z})$ lies in the x-z plane. Although the resulting physics is independent of gauge, this is a practical choice because this vector potential breaks translation symmetry in the y -direction only, meaning that the wave-functions remain plane-waves in the x and z directions.

The magnetic field enters our Dirac Hamiltonian, Eq. (7.1), by minimal coupling to the momentum operator $\hat{\mathbf{P}} \rightarrow \hat{\mathbf{P}} - \mathbf{A}$ resulting in the Hamiltonian (operator)

$$\hat{H} = \tau_z \left(\sigma_x \left(\hat{P}_x - B\hat{y} \pm v_F K_{0x} \right) + \sigma_y \hat{P}_y + \sigma_z (\hat{P}_z \pm v_F K_{0z}) \right). \quad (7.2)$$

In the same manner as for Schrödinger fermions, we can define ladder operators a and a^\dagger such that $a = \frac{1}{\sqrt{2}l_B B} (\hat{P}_x - v_F B\hat{y} - v_F K_{0x} + i\hat{P}_y)$ where $[\hat{a}, \hat{a}^\dagger] = 1$. In doing so we have established the quantity known as the **magnetic length**, $l_B = \sqrt{c/Be} = 26 \text{ nm}/\sqrt{\text{T}}$ (or simply $\sqrt{1/B}$ in our current units). Inserting this definition into the Hamiltonian operator we find that the off-diagonals of the σ matrices are complex conjugates and the full Hamiltonian can be written

$$\hat{H} = \tau_z \begin{pmatrix} \hat{P}_z \pm v_F K_{0z} & \Omega_0 \hat{a}^\dagger \\ \Omega_0 \hat{a} & -(\hat{P}_z \pm v_F K_{0z}) \end{pmatrix}, \quad (7.3)$$

where $\Omega_0 = v_F \sqrt{2}/l_B \sim 1/\sqrt{B}$ will determine the Landau level spacing.

Ignoring the \hat{P}_z term for the moment we see that this Hamiltonian squares to that of the Harmonic oscillator, which we know is solved by the number states $|n\rangle$ of the lowering operator a defined such that $a^\dagger a |n\rangle = n |n\rangle$. This leads us to propose the ansatz wave-function in terms of these number states, for $n \geq 1$ (we will discuss the special case of $n = 0$ separately):

$$|\psi_{k_z, n}; \pm \mathbf{K}_0\rangle = \begin{pmatrix} |n\rangle \\ |n-1\rangle \end{pmatrix} e^{i(k_z \pm K_{0z})z}. \quad (7.4)$$

Operating the Hamiltonian Eq. (7.3) on this state leads to an easy eigenvalue equation

$$\hat{H} |\psi_{k_z, n}, \tau; \pm \mathbf{K}_0\rangle = \tau_z \begin{pmatrix} v_F k_z & \Omega_0 \sqrt{n} \\ \Omega_0 \sqrt{n} & -v_F k_z \end{pmatrix} |\psi_{k_z, n}; \pm \mathbf{K}_0\rangle, \quad (7.5)$$

7.1. Landau levels in 3D Dirac materials

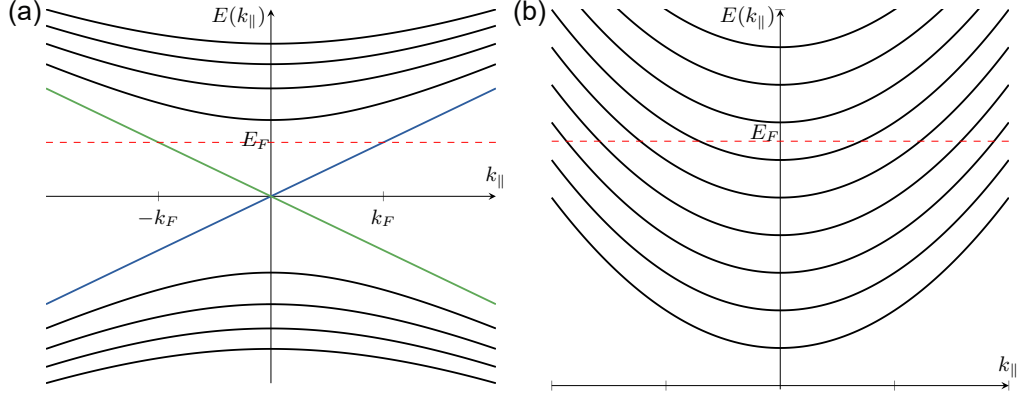


Fig. 7.1 **Landau levels of Dirac and Schrödinger materials in a magnetic field:** a) Landau levels of a Dirac material, the spacing between Landau levels decreases with energy and there are two gapless chiral modes in the lowest Landau level. Due to the level spacing in a Dirac material, it is possible to reach the lowest Landau level at relatively small magnetic field (E_F shown here is the location of the Fermi-level in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ at 14 T, see next chapter). b) The Landau levels of a Schrödinger material. Levels are equally spaced and there is a gap to zero energy. For the same magnetic field strength as (a) multiple Landau levels are occupied [109].

which has doubly degenerate eigenenergies arising from the τ_z matrix,

$$E_n(k_z, \tau = \pm; \pm \mathbf{K}_0) = \tau \sqrt{(v_F k_z)^2 + \Omega_0^2 n}. \quad (7.6)$$

We see that the application of a magnetic field leads to the quantisation of the electron's energy in terms of n . This is nothing but the well known Landau quantisation of a standard 3d electron gas described by the Schrödinger equation with each n resulting in (doubly degenerate) Landau level. What is, however, peculiar to Dirac fermions is that the spacing between Landau levels decreases as $\sim \sqrt{n} - \sqrt{n-1}$, whereas Schrödinger fermions have Landau levels that are equally spaced (see Fig. 7.1). The practical consequence of this is that the lower Landau levels can be reached at much lower magnetic fields and experiments more easily conducted deep in the lowest Landau level [110], $n = 0$, known as the “quantum limit”.¹

In the above discussion our wave-function was only defined for $n \geq 1$ and the $n = 0$ state avoided. As hinted by the name “quantum limit”, this state is actually the most interesting. Since $|n-1\rangle$ is not defined for $n = 0$ we need to slightly alter our ansatz for the lowest Landau level wave-function to

$$|\psi_{k_z,0}; \pm \mathbf{K}_0\rangle = \begin{pmatrix} |0\rangle \\ 0 \end{pmatrix} e^{i(k_z + K_{0z})z}. \quad (7.7)$$

¹A truly *practical* consequence is that such experiments can be done with $\sim 14\text{T}$ fields from superconducting magnetics found in many modern university laboratories and there is no need to resort to the much rarer and expensive resistive/superconducting-resistive magnetics required to produce larger fields.

Applying the Hamiltonian operator to this state yields only diagonal terms since $a|0\rangle = 0$. The eigenenergy can be read off as $E_0(k_z, n, \tau; \pm \mathbf{K}_0) = \tau k_z$ and is *singularly* degenerate, half the degeneracy of the other Landau levels. Strikingly this mode extends to zero-energy for $k_z = 0$ so will always be occupied, regardless of the magnetic field strength. Finally we note that the mode is chiral with a different τ_z index for positive and negative velocities. As a result, if impurities (or other scatterers) respect the underlying symmetry defining τ_z , backscattering from k_z to $-k_z$ within a Dirac cone is forbidden.

The corresponding wave-functions in real space – which will be especially useful when calculating Friedel oscillations later in the chapter – are calculated in Appendix D. The procedure is not dissimilar from that of calculating the eigenfunctions of the Harmonic oscillator. The result for the lowest Landau level is

$$\psi_0(\mathbf{r}) = A_0 \tau_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-(k_x l_B - y/l_B)^2/2} e^{i(k_z \pm K_{0z})z} e^{i(k_x \pm K_{0x})x}, \quad (7.8)$$

with normalisation $A_0 = 1/(\sqrt{l_B \sqrt{\pi} L_x L_z})$. We explicitly chose a gauge that maintained translation symmetry in the x and z direction so unsurprisingly the wave-function in these directions remains a plane wave. In the y -direction, which had its translation symmetry broken by the field, we find that for a given k_x the wave-function in the lowest Landau level decays as a Gaussian centred around $y = k_x$.

For the higher Landau levels the n th LL is given by,

$$\psi_n(k_x; y) = \begin{pmatrix} A_n H_n \left(\frac{y}{l_B} - k_x l_B \right) \\ A_{n-1} H_{n-1} \left(\frac{y}{l_B} - k_x l_B \right) \end{pmatrix} \exp \left\{ -\frac{1}{2} \left(\frac{y}{l_B} - k_x l_B \right)^2 \right\} e^{i(k_x + k_{0x})x} e^{i(k_z + k_{0z})z}, \quad (7.9)$$

where $H_n(\xi)$ is the n th Hermite polynomial and the normalisation is given by $A_n = 1/\sqrt{2^n n! l_B \sqrt{\pi}}$. We see from the presence of the Hermite polynomials that as the Landau level index n decreases the wave-functions become less extended in real space, until in the lowest Landau level they are quasi-one dimensional with a length-scale set by l_B in the directions perpendicular to the field.

7.2 Disordered Landau levels

Having established the eigenstates of clean Landau levels, we now consider the impact of disorder. We already mentioned in chapter 2 that disorder plays a key role of broadening the Landau levels that results in experimentally observable quantum Hall plateaus. In this section we will see that the localisation of the wave-function perpendicular to a magnetic field dramatically changes the response of electrons to impurities. Of particular interest for the next chapter will be that the LDOS and corresponding charge density that builds up around a scatterer is effectively 1 dimensional. As we saw in chapter 3, there is no decay in oscillations of the LDOS in 1d and the Friedel oscillations decay slowly as $\sim 1/r$.

7.2. Disordered Landau levels

7.2.1 Disordered 2d Landau levels

To begin with let us consider the impact of disorder on a Landau level in 2d, for example, the surface state of a topological insulator. The reason for first considering the impact of disorder in 2d is that, even for a 3d system, each momentum k_{\parallel} parallel to the field is endowed with a degenerate 2d Landau level perpendicular to the field. We will see that this results in many of the features of the 2d Landau level carrying over to those in three dimensions. The wave-functions of the 2d Landau levels are identical to those above but with $z = 0$.

For the rest of this sub-section we consider a 2d Landau level from a single Dirac cone in a magnetic field, for instance as occurs on the surface of a topological insulator (neglecting the Zeeman term) or graphene with long range disorder [111]. The free Green's function of a 2d Landau level can be written as an operator

$$\hat{G}_0(\omega) = \sum_{n=-\infty}^{\infty} \frac{|\psi_n\rangle \langle \psi_n|}{\omega - E_n}, \quad (7.10)$$

where ψ_n is the spinor describing the n -th Landau level and $|\psi_n\rangle \langle \psi_n|$ acts as a projector onto the n -th Landau level.

The LDOS in the vicinity of a single impurity point-like impurity $\mathbf{U}(\mathbf{r}) = u_0 \delta(\mathbf{r}) \mathbb{1}$. As we saw in chapter 3, the change in the LDOS is given by

$$\begin{aligned} \delta\rho(\mathbf{r}, \omega) &= -\text{Im} \int d\mathbf{r}_1 \text{Tr} \left\{ G_0(\mathbf{r} - \mathbf{r}_1) T(\mathbf{r}_1) G_0(\mathbf{r}_1 - \mathbf{r}) \right\} \\ &= -u_0 \text{Im} \sum_n \int \frac{dq}{2\pi} \frac{dk}{2\pi} e^{iqx} \text{Tr} \left\{ \frac{\psi_n(k; y) \psi_n^\dagger(k; 0) \psi_n(k+q; 0) \psi_n^\dagger(k+q; y)}{(\omega - E_n + i\delta)^2} \right\}, \end{aligned} \quad (7.11)$$

where in the second line we have used that impurity scattering is elastic and that there is some background broadening of the Landau level, set by δ . The imaginary part of the denominator gives the only energy dependent factor $\sim \frac{2(\omega - E_n)\delta}{(\delta^2 + (\omega - E_n)^2)^2}$, which simply enforces that we must be close to a Landau level to see any significant change in the density of states [112–114]. The only spatial dependence comes from the wave-functions themselves, which in the lowest Landau is particularly easy to calculate since the wave-functions are Gaussian, such that

$$\delta\rho(\mathbf{r}, \omega \approx 0) \sim e^{-\frac{x^2 + y^2}{2l_B^2}}. \quad (7.12)$$

In particular, the oscillatory behaviour of the LDOS that we saw in chapter 3 has completely disappeared. The reason for this is because the length scale which was set by $1/k_F$ has been replaced by the magnetic length $1/l_B$. Later in this section we will be interested in the charge density that forms around an impurity in a 3d Dirac semi-metal, we will see there that the only oscillatory behaviour remaining will be parallel to the magnetic field.

Having seen the response of our Landau levels to a single impurity, let's consider a dilute density, n_{imp} , of point-like, random, and diagonal impurities. Disorder averaging leaves the full Green's function diagonal in spin-space and Landau index since Hermite

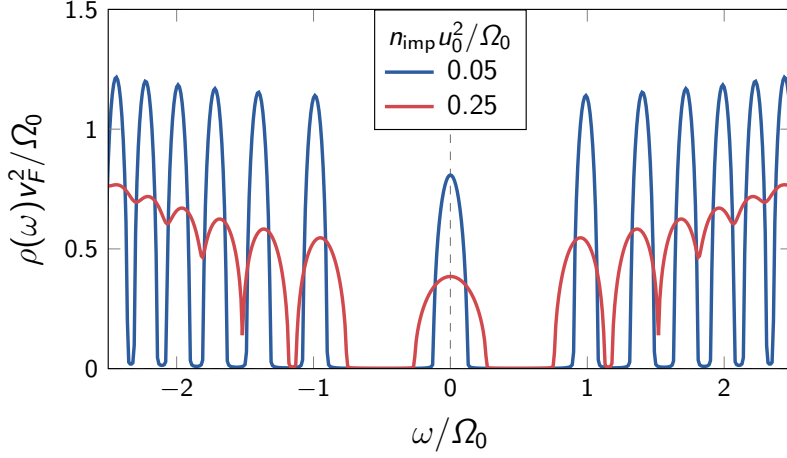


Fig. 7.2 Density of states of a 2d Dirac material in a magnetic field: The density of states of a single Dirac cone in a magnetic field, e.g. the surface of a topological insulator (excluding Zeeman contributions) calculated using the self-consistent Born approximation. For increasing energy ω Landau levels start to merge due to the smaller spacing between higher LLs, the merging occurs approximately when disorder strength $n_{\text{imp}} u_0^2$ and the Landau level spacing are equal [111].

polynomials of different orders n are orthogonal. The two diagonal components of the full disordered Green's function from the n -th Landau level $\mathcal{G}_n(\omega) = \text{diag}\{\mathcal{G}_n^+, \mathcal{G}_n^-\}$, are [111, 115]

$$\mathcal{G}_n^+ = \frac{\omega - \Sigma^-}{(\omega - \Sigma^+)(\omega - \Sigma^-) - \Omega_0^2 n} \quad (7.13)$$

$$\mathcal{G}_n^- = \frac{\omega - \Sigma^+}{(\omega - \Sigma^+)(\omega - \Sigma^-) - \Omega_0^2 (n+1)}, \quad (7.14)$$

where the self-energy, $\Sigma = \text{diag}\{\Sigma^+, \Sigma^-\}$, is no longer diagonal due to the matrix structure of the lowest Landau level. We also assume that these impurities are weak, so that the Born approximation is valid. In this limit the self-consistent Born approximation to the self-energy is

$$\Sigma(\omega) = \frac{n_{\text{imp}} u_0^2}{2\pi l_B^2} \sum_{n=0}^{N_c} \mathcal{G}_n(\omega), \quad (7.15)$$

where $\frac{1}{2\pi l_B^2}$ is the clean density of states per spin per Landau level, and N_c is an upper cut off of the highest Landau level.

As we have seen in previous chapters we can calculate the density of states using

$$\rho(\omega) = -\frac{1}{\pi} \sum_{k_x, n} \text{Tr}\{\text{Im } \mathcal{G}_n(\omega)\} = -\frac{1}{n_{\text{imp}} u_0^2 \pi} (\text{Im} \Sigma^+ + \text{Im} \Sigma^-). \quad (7.16)$$

The density of states calculated using the full self-consistent Born approximation calculated for two values of u_0^2 is shown in Fig. 7.2. The lowest Landau level is broader than the other Landau levels due to the Green's function only having a single component on the diagonal

7.2. Disordered Landau levels

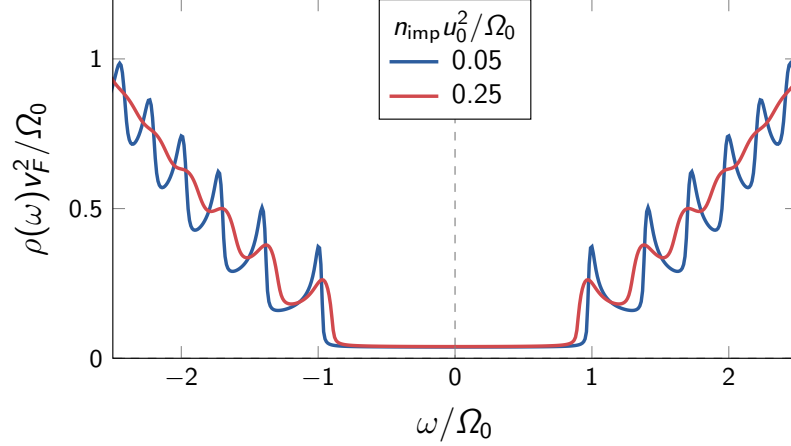


Fig. 7.3 **Density of states of a 3d Weyl semi-metal in a magnetic field:** The density of states of a single Weyl point in a magnetic field. Peaks in the density of state occur with increasing energy ω at the location of new Landau levels. As in the 2d case, Landau levels become ill defined for energies where the level spacing is of the order of the disorder strength [116].

corresponding to the energy of the lowest Landau level.² As energy increases the Landau levels begin to merge due to the decrease in spacing between them (see red line Fig. 7.2), this is very different to Schrödinger fermions since Landau levels are equally spaced and, if the lowest and 1st Landau level are well separated, all other Landau levels are separated from each other (within the self-consistent Born approximation).

7.2.2 Disordered 3d Landau levels

In an identical manner to the 2d Landau level, we can define the disorder averaged Green's function of a 3d Landau level as

$$\hat{G}_0(\omega, k_z) = \sum_{n=-\infty}^{\infty} \frac{|\psi_n\rangle \langle \psi_n|}{\omega - E_n(k_z)}, \quad (7.17)$$

which now contains a dependence on the momentum parallel to the magnetic field k_z .

Let us start by calculating the density of states and consider a single Weyl node located at the origin $\mathbf{K}_0 = 0$, of course we have seen in chapter 1 that there must be another Weyl node located at another point in the Brillouin zone, but for now we assume inter-nodal scattering can be neglected. The calculation is largely the same as in 2d but with an additional integral over k_z . The full disorder averaged Green's function for each Landau

²This is not the case for short-range scattering in graphene where inter-nodal scattering makes all Landau levels equally broad [111].

level n , $\mathcal{G}_n(\omega, k_z) = \text{diag}\{\mathcal{G}_n^+, \mathcal{G}_n^-\}$, is [116]

$$\mathcal{G}_n^+(k_z) = \frac{\omega - \Sigma^- + v_F k_z}{(\omega - \Sigma^+ - v_F k_z)(\omega - \Sigma^- + v_F k_z) - \Omega_0^2 n} \quad (7.18)$$

$$\mathcal{G}_n^-(k_z) = \frac{\omega - \Sigma^+ - v_F k_z}{(\omega - \Sigma^+ - v_F k_z)(\omega - \Sigma^- + v_F k_z) - \Omega_0^2(n+1)}. \quad (7.19)$$

Taking the same assumptions about impurities as in 2d we obtain the self-consistent self-energy equation in the Born limit,

$$\Sigma(\omega) = \frac{n_{\text{imp}} u_0^2}{2\pi l_B^2} \int \frac{dk_z}{2\pi} \mathcal{G}_n(\omega, k_z). \quad (7.20)$$

The equation for the density of states in terms of the self-energy is identical to the case in 2d (see Eq. (7.16)) and is shown for two values of impurity strength in Fig. 7.3. We see that, in a very similar manner to the topological insulator nanowires, the effective 1d physics of the Landau levels of a Weyl point give rise to oscillations in the density of states as new Landau levels become occupied. We will discuss these oscillations in the conductivity later in the chapter that related oscillations in the density of states as a function of magnetic field, known as “quantum oscillations”, result in the oscillatory behaviour of response functions, such as the conductivity.

7.2.3 Friedel oscillations in Dirac semi-metals in the quantum limit

To finish our discussion of disorder in Landau levels we calculate the change in the LDOS and resulting charge density that builds up around an impurity in a 3d Dirac semi-metal. We will now consider two Dirac points located at $\pm \mathbf{K}_0$ and write the free Green’s function

$$\hat{G}_0^{\xi, \chi}(\omega, k_z) = \sum_{n=-\infty}^{\infty} \frac{|\psi_n^{\xi, \chi}\rangle \langle \psi_n^{\xi, \chi}|}{\omega - E_n^{\chi}(k_z)}, \quad (7.21)$$

where $\xi = \pm$ labels the Dirac point and $\chi = \pm$ the chirality of the state. We focus on the lowest Landau level so the $E_0^{\chi}(k_z) = \chi v_F k_z$.

There are two potential back-scattering channels that will result in a significant change in charge density: 1) Intra-nodal scattering will strongly depend on the origin of the matrix structure of the Dirac Hamiltonian. In particular, if the impurities break the symmetries resulting in this matrix structure. For example, if the matrices arise due to a sub-lattice index and the impurities are local, then back-scattering will in general be allowed. 2) Inter-nodal scattering is not generically protected by any symmetry and so will result in a significant back-scattering channel unless the impurities are long-ranged so that the overlap between nodes is small (see next chapter). For the moment we will consider a single point-like impurity at the origin [117, 118]. The resulting local density of states in

7.2. Disordered Landau levels

the lowest Landau level is

$$\begin{aligned}\delta\rho^{\xi\xi}(\mathbf{r}, \omega) &= -\text{Im} \sum_{\chi\chi'} \int d\mathbf{r}_1 \text{Tr} \left\{ G_0^{\xi,\chi}(\mathbf{r} - \mathbf{r}_1) T(\mathbf{r}_1) G_0^{\xi,\chi'}(\mathbf{r}_1 - \mathbf{r}) \right\} \\ &= \text{Im} \sum_{\chi\chi'} \rho_0^{\chi\chi'} \int \frac{d\mathbf{q}}{(2\pi)^2} \frac{d\mathbf{k}}{(2\pi)^2} e^{i(q_x x + q_z z)} \frac{\psi_0(k_x; y) \psi_0^\dagger(k_x; 0) \psi_0(k_x + q_x; 0) \psi_0^\dagger(k_x + q_x; y)}{(\omega - v_F \chi k_z + i\delta)(\omega - v_F \chi'(k_z + q_z) + i\delta)},\end{aligned}\quad (7.22)$$

where $\rho_0^{\chi\chi'} \sim \langle \psi_0^{\xi,\chi} | \hat{U} | \psi_0^{\xi,\chi'} \rangle$ is a constant proportional to the overlap of the chiralities resulting from the impurity. Luckily we have already seen the perpendicular component in Eq. (7.11), which results in a Gaussian real space decay. This leaves only the quasi-1d k_z contribution parallel to the field

$$\delta\rho^{\xi\xi}(\mathbf{r}, \omega) = e^{-\frac{r_\perp^2}{4l_B^2}} \text{Im} \sum_{\chi\chi'} \rho_0^{\chi\chi'} \int \frac{dq_z}{2\pi} \frac{dk_z}{2\pi} e^{iq_z z} \frac{1}{(\omega - v_F \chi k_z + i\delta)(\omega - v_F \chi'(k_z + q_z) + i\delta)}.\quad (7.23)$$

The integral over k_z can be performed as follows: If $\chi = \chi'$, meaning that scattering is to a velocity v_F of the same sign, the poles of the denominator lie in the same half-plane and the integral is 0. This is as expected, since forward scattering does not cause any change in the LDOS. On the other hand if $\chi \neq \chi'$ and there is some overlap of the two chiralities,³ the integral can easily be solved by splitting into partial fractions and, after taking the imaginary part, we are left with an integral over a δ -function, such that

$$\delta\rho^{\xi\xi}(\mathbf{r}, \omega) = \tilde{\rho}_0 e^{-\frac{r_\perp^2}{4l_B^2}} \int \frac{dq_z}{2\pi} e^{iq_z z} \left(\frac{1}{2\omega + v_F q_z} - \frac{1}{2\omega - v_F q_z} \right) = \rho_0 e^{-\frac{r_\perp^2}{4l_B^2}} \sin\left(\frac{2\omega z}{v_F}\right),\quad (7.24)$$

where ρ_0 is a constant. This means the LDOS around an impurity is highly anisotropic: Parallel to the field the oscillations are those of a 1d system without any decay in real space.⁴ Perpendicular to the field the LDOS falls off as a Gaussian with a width set by the magnetic length l_B . This behaviour arises from the peculiar 1d nature of electrons in the lowest Landau level.

As we saw in chapter 3, the change in the LDOS also results in a change in the charge density, known as **Friedel oscillations**. These Friedel oscillations are simply the integral of the LDOS over energy, up until the Fermi-energy, such that

$$\delta n^{\xi\xi}(\mathbf{r}) = \int_{-\Lambda}^{\mu} d\omega \delta\rho^{\xi\xi}(\mathbf{r}, \omega) = \delta n_0 e^{-\frac{r_\perp^2}{4l_B^2}} \frac{\cos(2k_F z)}{|z|},\quad (7.25)$$

where $k_F = \mu/v_F$ is the Fermi-momentum and in the 2nd equality we have dropped the unphysical due to the hard cut-off Λ .

³If the impurities do not cause an overlap and back-scattering is forbidden we are in the same situation as chapter 3, now with two 1d chiral modes. There can be no singularities in the FT-LDOS but there could be divergences in derivatives which will lead to a faster decay of any oscillations.

⁴If we were to include dephasing effects, i.e. a finite self-energy, there will be a decay \sim the dephasing length.

Switching now to the contribution due to inter-nodal scattering between the two Dirac points located at $\pm \mathbf{K}_0$ the calculation of the LDOS and charge density is identical. The only modification comes from the products of wave-functions in Eq. (7.22), since there is a total phase difference $\Delta \mathbf{K}_0 = \pm 2\mathbf{K}_0 = (\Delta K_{0x}, 0, \Delta K_{0z})$ such that $\delta n^{\xi \neq \xi'}(\mathbf{r}) = e^{i\Delta \mathbf{K}_0 \cdot \mathbf{r}} \delta n^{\xi\xi}(\mathbf{r})$. In the next chapter we will see how interactions with these oscillations can result in a strong angular dependence of the magnetoresistance.

7.3 Transport of semi-metals in a magnetic field

Now that we have considered the impact of impurities on the density of states we switch to a brief discussion of transport properties of semi-metals in a magnetic field. The presence of Landau levels and their 1d nature for three dimensional systems leads to several fascinating phenomena, we will discuss just two in this section: The non conservation of chiral charge when electric and magnetic field are aligned, known as the chiral anomaly and quantum oscillations of the resistivity.⁵

7.3.1 Chiral anomaly

We know from Noether's theorem that continuous symmetries result in conservation laws. For instance the conservation of charge is a consequence of $U(1)$ invariance, this means that the system is invariant under multiplication of the wave-function by an arbitrary phase $\psi \rightarrow e^{i\theta} \psi$. We can associate the conserved quantity n to this conservation law, which in this case is, of course, the total charge. In chapter 1 we saw that, in the massless limit, the left and right handed modes of the Dirac equation decouple. This means that we can independently multiply the left or right-handed components of the wave-function by an arbitrary phase $\psi_L \rightarrow e^{i\theta_L} \psi'_L$ and $\psi_R \rightarrow e^{i\theta_R} \psi'_R$ [23]. In a similar manner to total charge we can associate the *chiral* charges n_L and n_R with these symmetries such that the net chiral charge $n_L - n_R$ should be conserved, if the system is massless. In the case of two Weyl nodes of opposite chirality this means that the charge on each node should be conserved.

We saw at the start of this chapter that applying a magnetic field to a Weyl node results in a single chiral channel in the lowest Landau level, with opposite Weyl nodes having opposite chirality. It is therefore clear that the conservation of chiral charge cannot be maintained in the presence of a magnetic field if there is a component of the electric field parallel to the magnetic field, this is because the electric field slightly reduces the chemical potential on one node and increases it on the other. In the presence of a term $\mathbf{E} \cdot \mathbf{B}$ the chiral charge is therefore *not* conserved. In particular the rate of change of chiral charge on each node can be written as

$$\frac{\partial(n_L - n_R)}{\partial t} = \frac{e^3}{2\pi\hbar^2} \mathbf{E} \cdot \mathbf{B}, \quad (7.26)$$

⁵Quantum oscillations are not unique to Dirac materials but can provide important information about the Fermi-surface that we will require in the next chapter.

7.3. Transport of semi-metals in a magnetic field

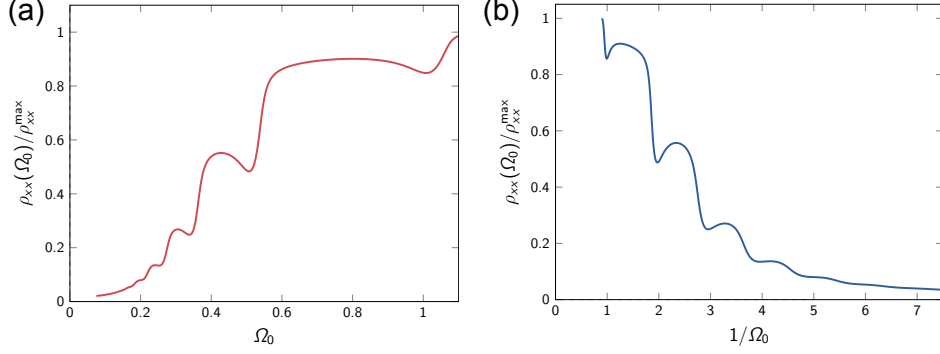


Fig. 7.4 **Quantum oscillations in the resistivity of a Weyl semi-metal:** a) The resistivity as a function of magnetic field in the self-consistent Born approximation using the same parameters as Fig. 7.3 (lowest impurity strength). b) The same plots as a function of the inverse field strength $1/\Omega_0$, using the inverse field strength, maxima are equally spaced.

of course in a real system such a term is eventually compensated by inter-nodal scatterings. Since the velocity of the charges on each node have opposite sign, this chiral charge pumping rate leads to a negative magnetoresistance [119], as we saw when discussing the Boltzmann equation in chapter 4. Such violations of conservation laws are known as anomalies. In this case the reason why the apparent conservation of chiral charge is violated stems from the fact that our low energy description of Weyl nodes ignored their connectivity in through the bulk [120].

7.3.2 Quantum oscillations in semi-metals

To calculate the conductivity we once again make use of the Kubo formula. As above we consider a single Weyl cone at the origin, $H = \boldsymbol{\sigma} \cdot \mathbf{p}$, such that the current operator is $\hat{\mathbf{j}} = ev_F \boldsymbol{\sigma}$. Ignoring vertex corrections and the negligible contributions from $\langle \mathcal{G} \mathcal{G} \rangle$ and $\langle \mathcal{G}^\dagger \mathcal{G}^\dagger \rangle$, the zero temperature transversal conductivity is [116, 121, 122]

$$\sigma_{xx}(B) = \frac{e^2 v_F^2}{2\pi l_B^2} \text{Re} \sum_{n=0}^{\infty} \int \frac{dk_z}{2\pi} \text{Tr} \text{Im} \mathcal{G}_n^+ \text{Im} \mathcal{G}_n^-, \quad (7.27)$$

where \mathcal{G}_n^\pm is the disorder averaged Green's function of the n -th Landau level in Eq. (7.19). A plot of the resistivity, calculated using the same SCBA result for the self-energy as in Fig. 7.3 is shown in Fig. 7.4a. We see that the oscillations in the density of states result in corresponding oscillations in the resistivity as a function of magnetic field. These oscillations are known as Shubnikov de Haas oscillations [123]. The frequency of these oscillations is inversely proportional to the magnetic field and set by the Fermi-energy and it is therefore possible to use these oscillations to measure the Fermi-wave vector (see e.g. discussion in [16]). We will see in the next chapter that this measurement will be enable us to predict the location of additional peaks in the experiment that we discuss there.

Summary and outlook

In this chapter we have seen that one of the unique properties of Dirac materials is their increased sensitivity to magnetic fields, with Landau level spacings proportional to \sqrt{n} . This enables us to more easily enter the quantum limit of the lowest Landau level. We also saw that disorder has the effect of broadening Landau levels and leads to quantum oscillations in the magnetoresistance.

Of particular importance for the next chapter is that, in the quantum limit, the Friedel oscillations that develop close to an impurity are effectively 1d in nature decaying in real space as $1/r$. In the next chapter we will see how this 1d physics can lead to a unique magnetoresistance phenomena deep in the quantum limit.

Transport spectroscopy of 3D Dirac materials in the quantum limit

In this chapter we turn our focus to a specific transport phenomenon in Dirac and Weyl semi-metals containing multiple Dirac or Weyl points. In the previous chapter we showed that a strong magnetic field \mathbf{B} causes electrons within each Dirac node to form effective one-dimensional channels flowing parallel or antiparallel to the field \mathbf{B} . The main idea behind this chapter is that the one-dimensional nature of electron motion deep in the quantum limit allows for scattering rates that can be controlled very efficiently via rotation of the magnetic field. More precisely, we will show that there is a strong angular dependence whenever there exists a scattering process that is sensitive to momentum transfer. This is because scattering between one dimensional channels is only dependent on the distance $|\Delta\mathbf{k}_{\parallel}|$, the projection of \mathbf{k} onto the direction parallel to the magnetic field. As a result magnetoresistance curves due to this effect can be used to “spectroscopically” map out the locations of electron pockets within the Brillouin zone.

So that we have a concrete experimental example, we begin this chapter by discussing the Dirac material $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ and angular dependent magnetoresistance experiments conducted upon it. After outlining theoretically the impact that momentum sensitive scattering has on magnetoresistance we turn to potential scattering mechanisms. The most obvious choice for a momentum dependent scatterer is a long-ranged impurity, however, we will show that such an impurity would need to be extremely broad to cause an effect as sharply angular dependent as that found in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$. In contrast to long range impurities, we will show that interactions between incoming electrons and the the quasi-1d Friedel oscillations, which we saw in the previous chapter, naturally result in extremely sharp angular dependence of magnetoresistance and predict well the location of the peaks in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$. While Friedel oscillations can reproduce the positions of peaks in magnetoreistance they struggle, at all but the largest fields, to reproduce the magnitude of the effect found in experiments.

8.1 Angular dependent magnetoresistance in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$

As discussed in chapter 1, Weyl nodes always come in pairs as monopoles and anti-monopoles connected by a quantized Berry flux and are only possible in materials that explicitly break either inversion or time reversal symmetry [23]. In contrast, it is in principle possible to have a material with a single Dirac node. Nonetheless, it is commonly the situation that further symmetries enforce the presence of multiple copies of Weyl- or Dirac nodes. For

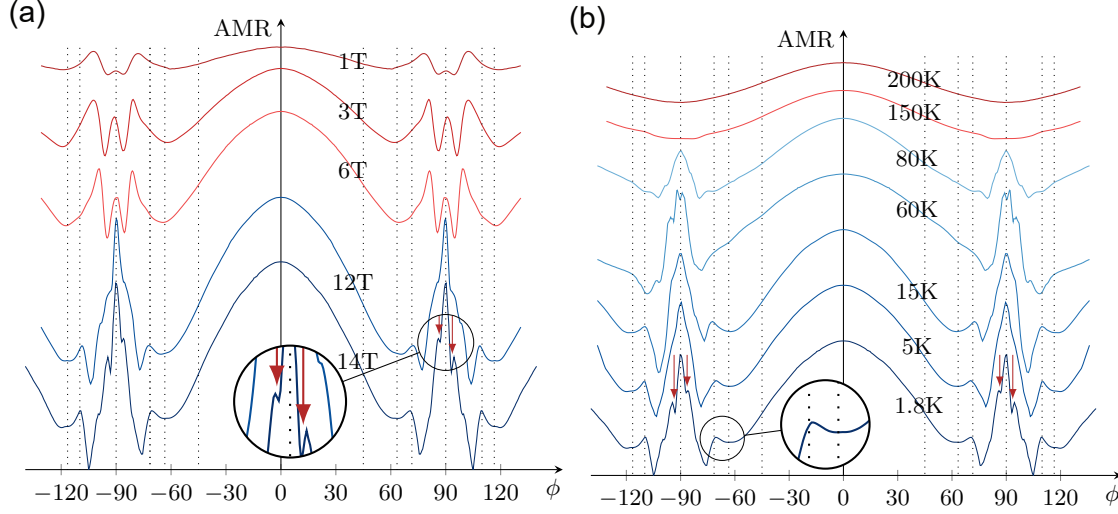


Fig. 8.1 **Anisotropic magnetoresistance in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$:** (a) The AMR measured at 1.8 K for a series of fixed magnetic fields. The data are symmetrized with respect to $+z$ and $-z$ direction to remove the admixture of Hall signal due to misalignment. Blue curves correspond to fields where the electrons primarily occupy the lowest Landau level, $B \sim 9\text{T}$ (b) AMR scans at 14 T for various temperatures. Blue curves indicate temperatures smaller than the spacing between the lowest and first Landau levels i.e. for these temperatures the vast majority of electrons occupy the lowest Landau level. In both panels the data are vertically shifted for clarity, dashed lines indicate the positions of magic angles corresponding to those in Fig. 8.2c, and arrows the position of extra peaks where the momentum difference between Fermi-pockets is $4k_F$, as predicted by our theory (see text).

example, TaAs [24] has 24 Weyl nodes and in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$, the Dirac semi-metal which the experiments discussed in this chapter are conducted on, there are four Dirac nodes [27].

We saw at the end of chapter 1 that $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ (PSS) undergoes a topological phase transition from a trivial insulator to a topological crystalline insulator as a function of x at low temperatures, the transition occurring at $x \approx 0.17$ [27]. At the phase transition the bulk band gap of PSS closes forming a Dirac semi-metal with 4 Fermi-pockets at the L-points of the bulk Brillouin zone (see Fig. 8.2a). The Hamiltonian at L-points was already presented in chapter 1, Eq. (1.27), the situation currently under consideration is $m \approx 0$.

The experiments discussed here were carried out on a low carrier concentration ($1.3 \times 10^{17} \text{ cm}^{-3}$ per pocket) single crystal of PSS with $x = 0.17$. The low electron density with correspondingly small Fermi momenta k_F and the large reciprocal space distance $\Delta k \sim 65k_F$ between Dirac nodes makes this system ideal to investigate how the momentum-space geometry affect transport.

The sample was mounted on a single-axis rotational probe so that the direction of the applied magnetic field can be rotated in the (010) plane with the current along the [100] axis. As the magnetic field is rotated there is a strong angular dependence of the angular

8.2. Geometry in a magnetic field

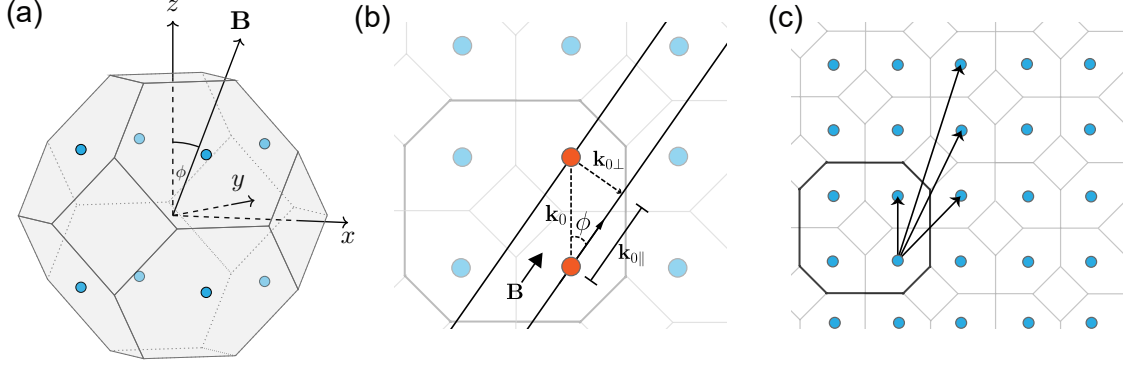


Fig. 8.2 **Geometry of $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ Brillouin zone:** a) The positions of the Fermi-pockets (blue) located at the L-points of the FCC Brillouin zone. Panels b) shows the geometry of this projection of the 3D Brillouin zone onto the xz plane, in which the magnetic field is rotated and c) Possible scattering vectors between different Fermi-pockets in this projected plane. At magic angles, the magnetic field is perpendicular to these scattering vectors (up to a small correction due to k_F , see e.g. Eq. (8.10)) giving rise to peaks in the resistivity.

magnetoresistance, $\text{AMR} \equiv (\rho_{xx}(\phi) - \rho_{xx}(\phi = 0)) / \rho_{xx}(\phi = 0)$, which is shown in Fig. 8.1 for a series of different magnetic fields (B) and temperatures (T). Around $\phi = 0^\circ$, the AMR shows the standard behaviour following $\sim \cos^2 \phi - 1$, which arises solely because the resistivity parallel to the magnetic field is smaller than perpendicular to the field. However, at around $\phi \simeq 70^\circ$, the AMR starts to deviate from the expected behaviour and AMR displays several pronounced peaks.

Surprisingly, the most pronounced peak occurs at $\phi = 90^\circ$ when the current flows parallel to the magnetic field. Naively we expect the lowest resistivity for this geometry since there is no orbital contribution and electrons can flow freely parallel to the field. The origin of the peaks cannot be ascribed to some peculiar Fermi surface shape, since the Shubnikov de Haas oscillations tell us that each Fermi surface is approximately spherical. The prominent angular dependence can therefore not originate from anisotropies of a single Fermi surface but is most likely related to the presence of multiple Fermi surfaces. Furthermore, Fig. 8.1b shows that the AMR changes only weakly as a function of temperature with the sharp peaks robust for a large range of temperature of 80 K down to the lowest measured temperatures, this rules out a phase transition [124] as a cause and strongly suggest that AMR is dominated by impurity scattering.

8.2 Geometry in a magnetic field

To understand the origin of the experimental peaks we need to consider the relevant geometry of the system after the application of the magnetic field. This is because, as we saw in the last chapter, parallel to the magnetic field the electrons can move freely, while perpendicular they perform a cyclotron motion. This means that in the quantum limit – where motion is essentially 1-dimensional – the Brillouin zone of the quasi-1d system in a

strong magnetic field is the *projection* of the Fermi-pockets onto the direction parallel to the magnetic field (see Fig. 8.2b). The projected distance in momentum space between two Fermi-pockets, $\mathbf{k}_{0\parallel}$, changes with the direction of the magnetic field as it is rotated in the (010) plane of $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$. There exist several “magic” angles, where the magnetic field is perpendicular to the momentum difference \mathbf{k}_0 between two Dirac nodes (shown in Fig. 8.2c) at these angles $\mathbf{k}_{0\parallel}$ vanishes. It is important to realise that these occur not just between Fermi-pockets within the same Brillouin zone but can also arise from scattering over several reciprocal lattice vectors.

Indeed for large magnetic fields and low temperatures, several (but not all) of the peaks occurring in the experimental AMR traces are at these magic angles where the projected inter-nodal momentum transfer, $\mathbf{k}_{0\parallel}$, vanishes (dashed lines in Fig. 8.1). This suggests the origin of the resistivity peaks is a scattering mechanism which can efficiently transfer momentum in the direction perpendicular to the field but is highly dependent on the momentum distance between Fermi-pockets parallel to the field.

Since peaks occur at locations where Fermi-pockets overlap one could use this measurement to *spectroscopically* map out the Brillouin zone of a Dirac or Weyl semi-metal if its bulk electronic structure was unknown.

8.3 Angular dependent scattering

For the scattering mechanism to be sensitive to the momentum transfer parallel to the field the scatterer responsible cannot be point-like, e.g. $U(\mathbf{r}) = u_0\delta(\mathbf{r} - \mathbf{r}_0)$, since the Fourier transform of such an impurity potential is a constant it scatters equally to all momenta. This sets a strong constraint on the underlying scattering mechanism responsible for the angular dependence of the magnetoresistance in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$; the scatterer must have a real space structure that is capable of inter-nodal scattering freely perpendicular to the field but not parallel.

8.3.1 Long range impurities

The most obvious candidate is a long ranged Gaussian impurity of width a with the form $U_{\text{imp}}^L(\mathbf{r}) = U_0^L e^{-(\mathbf{r}-\mathbf{r}_i)^2/2a^2}$, where the impurity is centered at \mathbf{r}_i . Scattering from such an impurity approximately conserves momentum. Since only the projected momentum parallel to the field is relevant very deep in the quantum limit, the idea is that a the Gaussian potential $U_{\text{imp}}^L(\mathbf{r})$ is capable of scattering between quasi-1d channels only when the Fermi-pockets are close in the projected momentum space, i.e. scattering is most efficient close to the magic angles discussed above.

As we saw in previous chapters, the scattering rate can be computed explicitly by calculating the inter-nodal matrix elements of the Landau level wave-functions with the

8.3. Angular dependent scattering

long-range potential, as appears in the Boltzmann equation. This gives

$$\Gamma_{k_x, k_x}^{k_z, k'_z} = \left\langle |\langle \psi_{k_x, k_z}^+ | U | \psi_{k_x, k'_z}^- \rangle|^2 \right\rangle_{\text{imp}} = M_0 e^{-(k'_x - k_x)^2 l_B^2 / 2} e^{-a^2 (k'_x - k_x + k_{0x})^2} e^{-a^2 (k'_z - k_z + k_{0z})^2}, \quad (8.1)$$

where \pm indicates states from the Dirac cone at $\pm \mathbf{k}_0/2$ and we have averaged over the impurity position \mathbf{r}_i . From the matrix element $\Gamma_{k_x, k_x}^{k_z, k'_z}$ we see that, in the true quantum limit¹ $l_B \rightarrow 0$, the factor $e^{-(k'_x - k_x)^2 l_B^2 / 2} \rightarrow 1$ and no longer suppresses large momenta transfers perpendicular to the field. In contrast, parallel to the field, the energy dependence of k_z is always pinned by the Fermi-energy such that $k_z, k'_z = \pm k_F$. This means that the transport scattering rate $\frac{1}{\tau_L^i}$ for scattering from Dirac node i to Dirac node j (possibly in a higher Brillouin zone) due to long-ranged impurities is given by

$$\frac{1}{\tau_{tr}^{ji}} = \int \frac{dk_x}{2\pi} \frac{dk'_x}{2\pi} \Gamma_{k'_x, k_x}^{\pm k_F, \mp k_F} = \frac{1}{\tau_0^{ji}} e^{-a^2 (\Delta k_{\parallel}^{ji})^2 - \frac{a^2 l_B^2 (\Delta k_{\perp}^{ji})^2}{2a^2 + l_B^2}} \sim e^{-\frac{2a^4 k_0^2}{2a^2 + l_B^2} (\phi^{ji} - \phi_0)^2}, \quad (8.2)$$

where $k_0 = |\mathbf{k}_j^0 - \mathbf{k}_i^0|$, $\Delta k_{\parallel, \pm}^{ji} = k_0 \cos \phi^{ji} \pm 2k_F$ and $\Delta k_{\perp}^{ji} = k_0 \sin \phi^{ji}$ are the momentum transfers parallel and perpendicular to the magnetic field, see Fig. 8.2. ϕ^{ji} denotes the angle between $\mathbf{k}_j^0 - \mathbf{k}_i^0$ and the magnetic field and $\phi_0 = \pm \left(\frac{\pi}{2} \pm 2k_F \frac{2a^2 + l_B^2}{2a^2 k_0} \right)$ determines the position of the peaks. Note that we only considered transitions between modes with *opposite* velocity, and therefore $|k_z - k'_z| = 2k_F$ because only those contribute to the resistivity and are relevant for transport scattering time.

From Eq. (8.2) we see that: (i) There is a peak in the scattering rate at “magic angles” whenever $\phi^{ji} \approx \phi_0$. The scattering vector \mathbf{k}_j^0 can be located in a higher-order Brillouin zone (BZ) (for \mathbf{k}_i^0 in the 1. BZ), suggesting that a sequence of many such peaks should exist at the angles defined in Fig. 8.2c. (ii) The position of peaks at these magic angles can be used to infer the position of the Dirac (or Weyl) nodes in the BZ.

Whilst the scattering rate due to an extended impurity is strongly angular dependent, in the final approximation of Eq. (8.2) we neglected a factor $\sim \exp\left(-\frac{l_B^2 k_0^2}{2a^2 + l_B^2}\right)$. This factor is unity in the quantum limit $l_B \rightarrow 0$ when the channels parallel to the field become truly one dimensional. Away from this limit, however, the overall size of the effect will be small unless the magnetic length is smaller than the extent of the impurity $l_B \ll a$ and the distance between Dirac points is small relative to the magnetic length $k_0 l_B \lesssim 1$. Both conditions are only fulfilled for very large fields. However, in reality, even for the largest field strengths used in the experiments 14 T, the magnetic length is a large length scale, in this corresponds to $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ the magnetic length is ~ 11 lattice spacings. As such it is unlikely these conditions are fulfilled for realistic field strengths because: 1) the size of l_B means that the first condition requires poor screening and 2) for the second condition

¹In this limit the Landau level associated with each k_z becomes extremely degenerate, such that there is a question of whether the material hosts a constant particle density, resulting in a B dependent μ , or a constant chemical potential, resulting in a B dependent n . When talking of this limit we will assume a constant chemical potential, this means there is a reservoir of charge that the system can use to fill up the LLs to maintain μ .

to be satisfied the Dirac points would need to be very close to each other in momentum space, this is not the case in our experimental system $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ where the Dirac points are located at the edges of the Brillouin zone, see Fig. 8.2.

8.3.2 Interaction corrections

There is, however, another scattering mechanism which can lead to highly anisotropic scatterer in our system. We saw in the previous chapter that internodal scattering from point-like impurities results in Friedel oscillations of the form

$$\delta n^{\xi\xi'}(\mathbf{r}) = n_0 e^{i\mathbf{k}_0 \cdot \mathbf{r}} e^{-\mathbf{r}_\perp^2/2l_B^2} \frac{\cos(2k_F z)}{|z|} = n_0 \delta n_\perp^{\xi\xi'}(\mathbf{r}_\perp) \delta n_\parallel^{\xi\xi'}(z), \quad (8.3)$$

where in the final equality we have used that the Friedel oscillations are separable. Consider now the interactions an electron with this Friedel oscillation, which we take to be a contact interaction $V_0(\mathbf{r} - \mathbf{r}') = V\delta(\mathbf{r} - \mathbf{r}')$. The Friedel oscillation then acts as a scatterer itself [125], such that

$$\delta V(\mathbf{r}_1, \mathbf{r}_2) = V_H(\mathbf{r}_1) \delta(\mathbf{r}_1 - \mathbf{r}_2) - V_F(\mathbf{r}_1, \mathbf{r}_2) \quad (8.4)$$

where we have defined the Hartree term

$$V_H(\mathbf{r}) = \int d\mathbf{r}' V_0(\mathbf{r}') \delta n(\mathbf{r} - \mathbf{r}') = V \delta n(\mathbf{r}) \quad (8.5)$$

and Fock term

$$V_F(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2} V_0(\mathbf{r}_1 - \mathbf{r}_2) \delta n(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2} V \delta n(\mathbf{r}_1) \delta(\mathbf{r}_1 - \mathbf{r}_2), \quad (8.6)$$

with $\delta n(\mathbf{r}_1, \mathbf{r}_2)$ the density matrix. The effective potential of the Friedel oscillation is therefore equivalent to an impurity of the form $u(\mathbf{r}) = V\delta n(\mathbf{r} - \mathbf{r}_i)$, where \mathbf{r}_i is the location of the impurity inducing the oscillation. These interaction corrections to the conductivity from 1d and 2d Friedel oscillations are well known [125–128] and are especially strong in low dimensions due to their slow decay.

In the case of inter-nodal Friedel oscillations the Fourier transform of the potential Eq. (8.3) is²

$$\delta n^{\xi\xi'}(\mathbf{q}) = \tilde{n}_0 e^{-(\mathbf{q}_\perp - \mathbf{k}_{0\perp})^2 l_B^2} \ln(|2k_F - q + k_{0\parallel}| |2k_F + q - k_{0\parallel}|), \quad (8.7)$$

²The arguments of the logarithms are made dimensionless by a term from the cut-off Λ that is approximately constant for small q (see previous chapter).

8.3. Angular dependent scattering

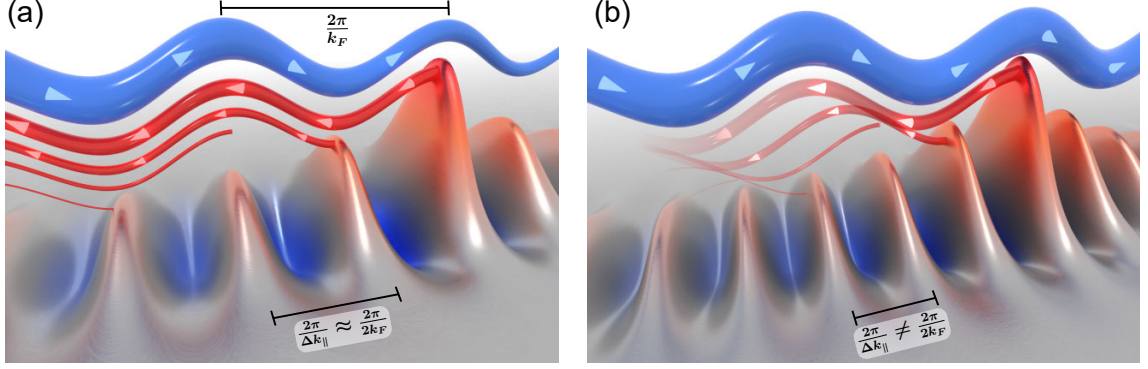


Fig. 8.3 **Scattering from Friedel oscillations:** Schematic image showing scattering of electrons from Friedel oscillations in the lowest Landau level (up to small wave length oscillations $\sim e^{ik_{0\perp}r_{\perp}}$). a) Close to magic angles the period of Friedel oscillations is $\approx \pi/k_F$ resulting in resonant back-scattering of incoming electrons. b) Away from magic angles the electrons are incoherently scattered and the effect is substantially smaller.

which we see in the quantum limit $l_B \rightarrow 0$ becomes truly one dimensional. The leading order correction to the self-energy is

$$\Sigma(k, \omega) = \left\langle \begin{array}{c} \text{Feynman diagram: a triangle with a dashed line on the left, a solid line on the right, and a wavy line at the bottom. The top vertex is marked with an 'X'. The left side is labeled with } \xi' \text{ and } \xi. \end{array} \right\rangle = u_0 V n_0 M_{\perp} \int \frac{dq}{2\pi} \frac{n_{\parallel}^{\xi\xi'}(k-q)}{\omega \pm v_F q}, \quad (8.8)$$

$$\text{where } M_{\perp} = \left\langle \int \frac{dq_x}{2\pi} \langle \psi_{k_x}^+ | n_{\perp} | \psi_{q_x}^- \rangle \langle \psi_{q_x}^+ | u_0 | \psi_{k_x}^- \rangle \right\rangle = M_0 e^{-l_B^2 k_{0\perp}^2 / 4} \quad (8.9)$$

is the impurity averaged matrix element of the perpendicular components, which is simply the matrix elements of the Landau wave functions with a Gaussian potential with width $a = l_B$ (see above). This correction represents an incoming electron scattering off first the Friedel oscillation and then impurity center itself. This contribution is non-zero after impurity averaging because the center of the Friedel oscillation is the impurity which is scattered off, such that $\langle \delta n(\mathbf{q}) u(\mathbf{q}) \rangle_{\text{imp}} \neq 0$. One may expect that the long range of the Friedel oscillations will result in a large amount of interference with other oscillations, however the only terms that survive impurity averaging parallel to the field are scatterings involving the same impurity, to linear order in impurity density n_{imp} .

From the self-energy we are able to calculate the scattering rate at the Fermi-momenta $k_F = \mu/v_F$, such that

$$\frac{1}{2\tau_{k_F}} = -\text{Im}\Sigma(\mu, k_F) = u_0 V n_0 M_{\perp} n_{\parallel}^{\xi\xi'}(k - k_F) = \frac{1}{2\tau_0} \ln(|4k_F - k_{0\parallel}| |k_{0\parallel}|). \quad (8.10)$$

We see that this rate is has a maxima at two positions, when the projection of the distance between Fermi-pockets $k_{0\parallel} = 0, 4k_F$. The equivalent Friedel oscillation due to the reverse inter-nodal scattering process will have maxima at $k_{0\parallel} = 0, -4k_F$. This means that, in

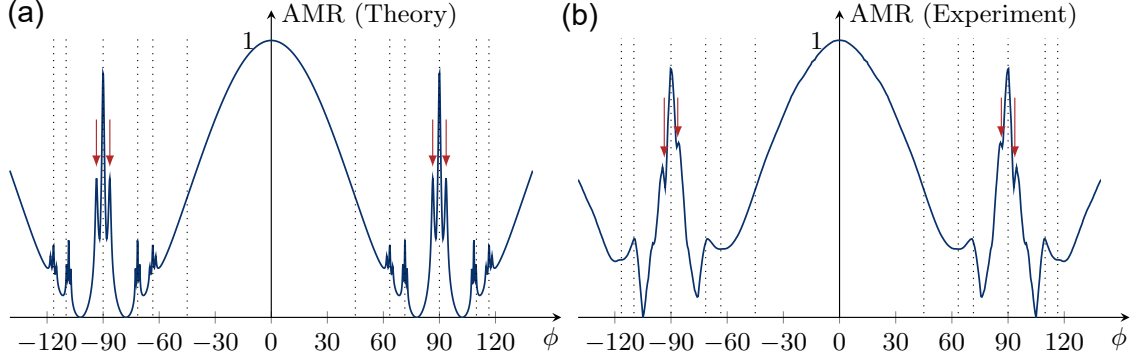


Fig. 8.4 **Comparison of AMR theory and experiment:** a) Plot of theory resistivity using the scattering time Eq. (8.10), in the quantum limit b) Experiment at 14 T and 1.8 K. Theory predicts the location of peaks in the resistivity at magic angles where the projected distance between Fermi-pockets vanishes. Additional peaks are expected when this distance $k_{0\parallel} = \pm 4k_F$, these are indicated by red arrows.

the quantum limit where $M_{\perp} \sim 1$, we expected extremely strong scattering that is highly dependent on angle. In addition to a peak when the field is perpendicular to the vector connecting two Fermi-pockets \mathbf{k}_0 , we also expect a second peak when this projected distance $k_{0\parallel} = \pm 4k_F$. Physically the origin of these peaks comes from angles where the overall wave-length of the inter-nodal Friedel oscillations is π/k_F , when this occurs the Friedel oscillation resonantly back-scatters incoming electrons (see Fig. 8.3a), away from these angles the Friedel oscillation has a period that results in scattering that is off-resonant (see Fig. 8.3b).

As we saw at the end of the previous chapter, we are able to independently measure k_F using the period of Shubnikov de Haas oscillations, such that it is found $k_F = 1.6 \times 10^6 \text{ cm}^{-1}$ which for the distance k_0 between the closest Fermi-pockets gives $k_0 \approx 65k_F$ our theory therefore predicts peaks at 86.4° and 93.6° . These are shown by red arrows in Fig. 8.4. We see that our theory therefore also explains the locations of the side peaks found by experiment.

Just as for the long-ranged impurities, however, the experiment is not conducted in the true quantum limit. This means that the matrix element M_{\perp} resulting from scattering intra-nodal scattering off the rapidly oscillating gaussian envelope perpendicular to the field, is extremely small when $l_B k_{0\perp} \gg 1$, as is the case even at the highest experimental fields. Unlike the long-ranged impurities, the Friedel oscillations can still result in a large scattering rate if there is an additional sharpening of this envelope function along the line close to $\mathbf{r}_{\perp} \approx 0$. We speculate that this sharpening may occur, for example, due to additional interactions of electrons with themselves or other impurities.

8.3.3 Resistivity in the quantum limit

As discussed in chapter 4, the simplest approximation to the conductivity is the Drude conductivity σ_D . From the definition of the scattering time and using the fact that the

8.3. Angular dependent scattering

density of states for each k_{\parallel} within a single Landau level is given by $1/2\pi l_B^2$ the conductivity due to either of the scatterers discussed above is given by

$$\sigma_{\parallel} = N_D N_S \frac{e^2 v_F \tau_{tr}(k_F)}{4\pi^2 \hbar l_B^2}, \quad (8.11)$$

where N_S and N_D are the number of spins per states and the number of Dirac nodes in the system, respectively. The resistivity using this formula calculated in the quantum limit and including the $\cos^2(\phi)$ dependence of ρ_{\perp} is plotted in Fig. 8.4a.

Summary and outlook

In this chapter we saw that $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ shows several pronounced resistivity maxima as a function of the angle of an applied magnetic field. The location of these peaks can be predicted by simple geometric arguments due to the presence of multiple Fermi-surface in the Brillouin zone. This is because only the projection of the Fermi-pockets onto the direction parallel to the field is important in the extreme quantum limit $l_B \rightarrow 0$. We also discussed two momentum dependent scattering mechanisms that could give rise to such maxima in resistivity: Long range impurities and the interactions of incoming electrons with Friedel oscillations that build up in the vicinity of an impurity. In particular we argued that, close to angles where the Fermi-pockets are located on top of each other in the projected space parallel to the field, scattering is approximately resonant. In contrast as the field is rotated away from these directions the scattering rate is significantly diminished. Not only does our theory predicts maxima close to these resonant “magic angles” but it is also capable of reproducing the locations of side-peaks which we predict to occur when Fermi-surfaces are separated by $4k_F$.

Surprisingly, the experimentally observed effect is significantly more enhanced than our theory predicts for even the maximum experimental field strength. This strongly suggests that either the Friedel oscillations or the Landau levels in $\text{Pb}_{1-x}\text{Sn}_x\text{Se}$ are more one dimensional than expected.

Conclusions and outlook

In this thesis we investigated the impact of disorder on Dirac materials with a particular focus on transport phenomena. As we have seen throughout, the ever-growing number of theoretical and experimental systems that are governed by a Dirac-like Hamiltonian provides a constant stream of novel phenomena, as well as the possibility of resolving many high energy questions in a condensed matter setting. This incredibly fast pace of development is evinced by the birth of topological insulators, which went from theoretical prediction to experimental realisation within just 2 years. Dirac materials not only provide a playground for novel quantum phenomena, but it is also likely that many future technological advances will utilise these materials.

An important feature of Dirac materials is their protection from back-scattering when disorder preserves the symmetries which protect their Dirac cone. We saw that one way to measure this is to look for the faster decay of oscillations in the local density of states that form close to an impurity. In particular, it is possible to experimentally observe these oscillations using scanning tunnelling spectroscopy on the surface of a topological insulator. By combining theory and experiment we were able to measure this protection and, in doing so, characterise the surface of the topological insulator.

Having established the robust back-scattering protection provided by a topological insulator's surface, we proceeded to break it by the application of an in-plane magnetic field. Such a field does not destroy the Dirac physics of the surface states but enables the scattering of electrons with a momentum perpendicular to the field. The result is an anisotropic magnetoresistance and an associated planar Hall effect which was also experimentally observed on the surface of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ thin films. From a theoretical perspective it was important to include the vertex corrections of the current-current correlation function, which encode the conservation laws necessary to correctly calculate transport properties.

When the surface of a topological insulator is wrapped up to form a thin nanowire, the surface states have a quantised angular momentum around the wire. As a result the kinetic energy along the wire is governed by bands with an equal energy spacing. We showed that these energy sub-bands result in a series of peaks in the resistivity of a topological insulator nanowire as a function of chemical potential. By comparing our theory with experiments on very thin wires of $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ we were able to unambiguously identify the confinement of the wave-function around the wire.

Finally, we saw that 3d Dirac materials in a magnetic field become effectively one dimensional parallel to the field. This is due to the reduction in spacing between Landau levels in a Dirac material, which enables the system to reach the quantum limit of the lowest Landau level even for relatively small fields. In this quantum limit the presence of multiple Fermi-surface in the Brillouin zone, coupled with the effectively 1d Friedel oscillations that build up around impurities, result in a highly angular dependent magnetoresistance. This occurred because only the projection of the Fermi-pockets onto the direction parallel to the field is important in the quantum limit. The highly momentum dependent scattering that results from interactions of electrons with Friedel oscillations causes an extremely strong resonant back-scattering close to angles where the Fermi-pockets are located on top of each other in this projected space and is significantly diminished away from these angles. Our theory predicts well the locations of peaks in the magnetoresistance seen in $\text{Pb}_x\text{Sn}_{1-x}\text{Se}$. Puzzlingly, however, the experimentally observed effect is significantly more enhanced than our theory expects at the field strengths used in real experiments.

Outlook: At the end of each chapter we have provided a more detailed outlook, here we will only discuss some of the most exciting prospects for future work:

The controlled breaking of time-reversal symmetry provides a mechanism to lift the topological protection of a topological insulator. It is interesting to consider further impacts this might have on the surface state, for example, whether it is possible to induce localisation using this effect. The lifting of protection may also prove important for nanowires, where the threading of half a flux quantum along the wire leads to the closing of the surface band gap.

Topological insulator nanowires present one of the most exciting frontiers for novel physics in Dirac materials. Their relative infancy leaves many unanswered theoretical and experimental questions unanswered. The most obvious experimental goal is the realisation of Majorana zero modes at the ends of a wire and the potential manipulation of these modes to perform quantum computation. In the more immediate future a detailed theoretical understanding of nanowire's transport properties in the presence of a magnetic field is required, including effects such as the magnetised impurities that induced the planar Hall effect. Also the effects of geometry and charged impurities, which may induce puddles in the bulk or on the surface of the nanowire, are not yet understood.

The excellent fit of theoretical and experimental peaks in $\text{Pb}_x\text{Sn}_{1-x}\text{Se}$ indicates that the electronic states and/or Friedel oscillations of 3d Dirac materials are more one-dimensional than we expect. This clearly calls for an explanation as to why the effect is so strong in $\text{Pb}_x\text{Sn}_{1-x}\text{Se}$ and whether a similar such effect occurs in other Dirac semi-metals or Weyl semi-metals.

Of course the most exciting frontier is not the questions we already know about, but those experimental and theoretical surprises that lie around the corner, waiting to be discovered.

References

- [1] Novoselov, K. S. *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197 (2005).
- [2] Wehling, T., Black-Schaffer, A. M. & Balatsky, A. V. Dirac materials. *Adv. Phys.* **63**, 1–76 (2014).
- [3] Nielsen, H. B. & Ninomiya, M. The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal. *Phys. Lett. B* **130**, 389–396 (1983).
- [4] Huang, X. *et al.* Observation of the chiral-anomaly-induced negative magnetoresistance in 3d Weyl semimetal TaAs. *Phys. Rev. X* **5**, 031023 (2015).
- [5] Geim, A. K. & Novoselov, K. S. The rise of graphene. In *Nanoscience and Technology: A Collection of Reviews from Nature Journals*, 11–19 (World Scientific, 2010).
- [6] Sarma, S. D., Adam, S., Hwang, E. & Rossi, E. Electronic transport in two-dimensional graphene. *Rev. Mod. Phys.* **83**, 407 (2011).
- [7] Haldane, F. D. M. Model for a quantum Hall effect without landau levels: Condensed-matter realization of the " parity anomaly". *Phys. Rev. Lett.* **61**, 2015 (1988).
- [8] Kane, C. L. & Mele, E. J. Quantum spin Hall effect in graphene. *Phys. Rev. Lett.* **95**, 226801 (2005).
- [9] Kane, C. L. & Mele, E. J. Z₂ topological order and the quantum spin Hall effect. *Phys Rev Lett* **95**, 146802 (2005).
- [10] Bernevig, B. A., Hughes, T. L. & Zhang, S.-C. Quantum spin Hall effect and topological phase transition in hgte quantum wells. *Science* **314**, 1757–1761 (2006).
- [11] Fu, L., Kane, C. L. & Mele, E. J. Topological insulators in three dimensions. *Phys. Rev. Lett.* **98**, 106803 (2007).
- [12] Moore, J. E. & Balents, L. Topological invariants of time-reversal-invariant band structures. *Phys. Rev. B* **75**, 121306 (2007).
- [13] Roy, R. Topological phases and the quantum spin Hall effect in three dimensions. *Phys. Rev. B* **79**, 195322 (2009).
- [14] Hasan, M. Z. & Kane, C. L. Colloquium: topological insulators. *Rev. Mod. Phys.* **82**, 3045 (2010).
- [15] Qi, X.-L. & Zhang, S.-C. Topological insulators and superconductors. *Rev. Mod. Phys.* **83**, 1057 (2011).
- [16] Ando, Y. Topological insulator materials. *J. Phys. Soc. Jpn.* **82**, 102001 (2013).

References

- [17] König, M. *et al.* Quantum spin Hall insulator state in HgTe quantum wells. *Science* **318**, 766–770 (2007).
- [18] Hsieh, D. *et al.* A topological Dirac insulator in a quantum spin Hall phase. *Nature* **452**, 970 (2008).
- [19] Hsieh, D. *et al.* Observation of unconventional quantum spin textures in topological insulators. *Science* **323**, 919–922 (2009).
- [20] Burkov, A. & Balents, L. Weyl semimetal in a topological insulator multilayer. *Phys. Rev. Lett.* **107**, 127205 (2011).
- [21] Young, S. M. *et al.* Dirac semimetal in three dimensions. *Phys. Rev. Lett.* **108**, 140405 (2012).
- [22] Cayssol, J. Introduction to Dirac materials and topological insulators. *C. R. Phys.* **14**, 760–778 (2013).
- [23] Armitage, N., Mele, E. & Vishwanath, A. Weyl and Dirac semimetals in three-dimensional solids. *Rev. Mod. Phys.* **90**, 015001 (2018).
- [24] Lv, B. *et al.* Experimental discovery of Weyl semimetal TaAs. *Phys. Rev. X* **5**, 031013 (2015).
- [25] Xu, S.-Y. *et al.* Discovery of a Weyl fermion semimetal and topological fermi arcs. *Science* **349**, 613–617 (2015).
- [26] Hsieh, T. H. *et al.* Topological crystalline insulators in the snt material class. *Nat. Commun.* **3**, 982 (2012).
- [27] Ando, Y. & Fu, L. Topological crystalline insulators and topological superconductors: from concepts to materials. *Annu. Rev. Condens. Matter Phys.* **6**, 361–381 (2015).
- [28] Tsuei, C. & Kirtley, J. Pairing symmetry in cuprate superconductors. *Rev. Mod. Phys.* **72**, 969 (2000).
- [29] Neto, A. C., Guinea, F., Peres, N. M., Novoselov, K. S. & Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **81**, 109 (2009).
- [30] Kitaev, A. Anyons in an exactly solved model and beyond. *Ann. Phys.* **321**, 2–111 (2006).
- [31] O’Brien, K., Hermanns, M. & Trebst, S. Classification of gapless Z2 spin liquids in three-dimensional kitaev models. *Phys. Rev. B* **93**, 085101 (2016).
- [32] Hermanns, M., Kimchi, I. & Knolle, J. Physics of the kitaev model: Fractionalization, dynamic correlations, and material connections. *Annu. Rev. Condens. Matter Phys.* **9**, 17–33 (2018).
- [33] Dirac, P. A. M. The quantum theory of the electron. *Proc. Roy. Soc.* **117**, 610–624 (1928).
- [34] Dirac, P. A. M. A theory of electrons and protons. *Proc. Roy. Soc.* **126**, 360–365 (1930).
- [35] Anderson, C. D. The positive electron. *Phys. Rev.* **43**, 491 (1933).

- [36] Zhao, J. *et al.* Rise of silicene: A competitive 2d material. *Prog. Mater. Sci.* **83**, 24–151 (2016).
- [37] Acun, A. *et al.* Germanene: the germanium analogue of graphene. *J Phys : Condens Matter* **27**, 443002 (2015).
- [38] Khandelwal, A., Mani, K., Karigerasi, M. H. & Lahiri, I. Phosphorene—the two-dimensional black phosphorous: Properties, synthesis and applications. *Materials Science and Engineering: B* **221**, 17–34 (2017).
- [39] Jotzu, G. *et al.* Experimental realization of the topological haldane model with ultracold fermions. *Nature* **515**, 237 (2014).
- [40] Semenoff, G. W. Condensed-matter simulation of a three-dimensional anomaly. *Phys. Rev. Lett.* **53**, 2449 (1984).
- [41] Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. Roy. Soc.* **392**, 45–57 (1984).
- [42] Simon, B. Holonomy, the quantum adiabatic theorem, and Berry’s phase. *Phys. Rev. Lett.* **51**, 2167 (1983).
- [43] Pancharatnam, S. Generalized theory of interference and its applications. In *Proceedings of the Indian Academy of Sciences-Section A*, vol. 44, 398–417 (Springer, 1956).
- [44] Wilczek, F. & Shapere, A. *Geometric phases in physics*, vol. 5 (World Scientific, 1989).
- [45] Sichau, J. *et al.* Resonance microwave measurements of an intrinsic spin-orbit coupling gap in graphene: a possible indication of a topological state. *Phys. Rev. Lett.* **122**, 046403 (2019).
- [46] Fu, L. & Kane, C. L. Time reversal polarization and a z^2 adiabatic spin pump. *Phys. Rev. B* **74**, 195312 (2006).
- [47] Drozdov, I. K. *et al.* One-dimensional topological edge states of bismuth bilayers. *Nat. Phys.* **10**, 664 (2014).
- [48] Liu, C., Hughes, T. L., Qi, X.-L., Wang, K. & Zhang, S.-C. Quantum spin Hall effect in inverted type-II semiconductors. *Phys. Rev. Lett.* **100**, 236601 (2008).
- [49] Knez, I., Du, R.-R. & Sullivan, G. Evidence for helical edge modes in inverted InAs/GaSb quantum wells. *Phys. Rev. Lett.* **107**, 136603 (2011).
- [50] Ostrovsky, P., Gornyi, I. & Mirlin, A. Quantum criticality and minimal conductivity in graphene with long-range disorder. *Phys. Rev. Lett.* **98**, 256801 (2007).
- [51] Fu, L. Hexagonal warping effects in the surface states of the topological insulator Bi₂Te₃. *Phys. Rev. Lett.* **103**, 266801 (2009).
- [52] Alpichshev, Z. *et al.* STM imaging of electronic waves on the surface of Bi₂Te₃: topologically protected surface states and hexagonal warping effects. *Phys. Rev. Lett.* **104**, 016401 (2010).

References

- [53] Chen, Y. *et al.* Massive Dirac fermion on the surface of a magnetically doped topological insulator. *Science* **329**, 659–662 (2010).
- [54] Taskin, A. A., Ren, Z., Sasaki, S., Segawa, K. & Ando, Y. Observation of Dirac Holes and Electrons in a Topological Insulator. *Phys. Rev. Lett.* **107**, 016801–016804 (2011).
- [55] Xu, N. *et al.* Surface and bulk electronic structure of the strongly correlated system SmB6 and implications for a topological Kondo insulator. *Phys. Rev. B* **88**, 121102 (2013).
- [56] von Neuman, J. & Wigner, E. Über merkwürdige diskrete eigenwerte. über das verhalten von eigenwerten bei adiabatischen prozessen. *Phys. Z.* **30**, 467–470 (1929).
- [57] Turner, A. M. & Vishwanath, A. Beyond band insulators: topology of semimetals and interacting phases. *Book chapter* **6**, 293–324 (2013).
- [58] Nielsen, H. B. & Ninomiya, M. Absence of neutrinos on a lattice:(ii). intuitive topological proof. *Nucl. Phys. B* **193**, 173–194 (1981).
- [59] Liu, Z. *et al.* Discovery of a three-dimensional topological Dirac semimetal, Na₃Bi. *Science* **343**, 864–867 (2014).
- [60] Balatsky, A. V., Vekhter, I. & Zhu, J.-X. Impurity-induced states in conventional and unconventional superconductors. *Rev. Mod. Phys.* **78**, 373 (2006).
- [61] Hermanns, M., O’Brien, K. & Trebst, S. Weyl spin liquids. *Phys. Rev. Lett.* **114**, 157202 (2015).
- [62] Coleman, P. *Introduction to Many-Body Physics* (Cambridge University Press, 2015).
- [63] Klitzing, K. v., Dorda, G. & Pepper, M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.* **45**, 494 (1980).
- [64] Tong, D. Lectures on the quantum Hall effect. *arXiv preprint arXiv:1606.06687* (2016).
- [65] Ostrovsky, P., Gornyi, I. & Mirlin, A. Electron transport in disordered graphene. *Phys. Rev. B* **74**, 235443 (2006).
- [66] Mahan, G. D. *Many Particle Physics* (Plenum, New York, 2000).
- [67] Bruus, H. & Flensberg, K. *Many-body quantum theory in condensed matter physics: an introduction* (Oxford university press, 2004).
- [68] Kohn, W. & Luttinger, J. Quantum theory of electrical transport phenomena. *Phys. Rev.* **108**, 590–611 (1957).
- [69] Altland, A. & Simons, B. D. *Condensed matter field theory* (Cambridge university press, 2010).
- [70] Taskin, A. *et al.* Planar Hall effect from the surface of topological insulators. *Nat. Commun.* **8**, 1340 (2017).
- [71] Friedel, J. Metallic alloys. *Il Nuovo Cimento (1955-1965)* **7**, 287–311 (1958).

- [72] Bena, C. Friedel oscillations: Decoding the hidden physics. *C. R. Phys.* **17**, 302–321 (2016).
- [73] Lee, W.-C., Wu, C., Arovas, D. P. & Zhang, S.-C. Quasiparticle interference on the surface of the topological insulator Bi₂Te₃. *Phys. Rev. B* **80**, 245439 (2009).
- [74] Pereg-Barnea, T. & MacDonald, A. H. Chiral quasiparticle local density of states maps in graphene. *Phys. Rev. B* **78**, 014201 (2008).
- [75] Guo, H.-M. & Franz, M. Theory of quasiparticle interference on the surface of a strong topological insulator. *Phys. Rev. B* **81**, 041102 (2010).
- [76] Bena, C. Green’s functions and impurity scattering in graphene. *Phys. Rev. B* **79**, 125427 (2009).
- [77] Knispel, T. *et al.* Charge puddles in the bulk and on the surface of the topological insulator bisbte_{se}₂ studied by scanning tunneling microscopy and optical spectroscopy. *Phys. Rev. B* **96**, 195135 (2017).
- [78] Bömerich, T., Lux, J., Feng, Q. T. & Rosch, A. Length scale of puddle formation in compensation-doped semiconductors and topological insulators. *Phys. Rev. B* **96**, 075204 (2017).
- [79] Kim, S. *et al.* Robust protection from backscattering in the topological insulator Bi_{1.5}Sb_{0.5}Te_{1.7}Se_{1.3}. *Phys. Rev. Lett.* **112**, 136802 (2014).
- [80] Ziman, J. *Electrons and Phonons: The Theory of Transport Phenomena in Solids*. International series of monographs on physics (OUP Oxford, 2001).
- [81] Carpentier, D. *Transport of Dirac Surface States* (2014).
- [82] Arakane, T. *et al.* Tunable Dirac cone in the topological insulator Bi(2-x)SbxTe(3-y)Sey. *Nat. Commun.* **3**, 636 (2012).
- [83] Sze, S. M. & Ng, K. K. *Physics of semiconductor devices* (John wiley & sons, 2006).
- [84] Borgwardt, N. *et al.* Self-organized charge puddles in a three-dimensional topological material. *Phys. Rev. B* **93**, 245149 (2016).
- [85] Breunig, O. *et al.* Gigantic negative magnetoresistance in the bulk of a disordered topological insulator. *Nat. Commun.* **8**, 15545 (2017).
- [86] Goldman, A. Electrostatic gating of ultrathin films. *Annu. Rev. Mater. Res.* **44**, 45–63 (2014).
- [87] Xia, J., Chen, F., Li, J. & Tao, N. Measurement of the quantum capacitance of graphene. *Nat. Nanotech.* **4**, 505 (2009).
- [88] Yang, F. *et al.* Dual-gated topological insulator thin-film device for efficient fermi-level tuning. *ACS Nano* **9**, 4050–4055 (2015).
- [89] Zyuzin, A., Hook, M. & Burkov, A. Parallel magnetic field driven quantum phase transition in a thin topological insulator film. *Phys. Rev. B* **83**, 245428 (2011).
- [90] Zyuzin, A. & Burkov, A. Thin topological insulator film in a perpendicular magnetic field. *Phys. Rev. B* **83**, 195413 (2011).

References

- [91] Biswas, R. R. & Balatsky, A. V. Impurity-induced states on the surface of three-dimensional topological insulators. *Phys. Rev. B* **81**, 233405 (2010).
- [92] Tang, H., Kawakami, R., Awschalom, D. & Roukes, M. Giant planar Hall effect in epitaxial (ga, mn) as devices. *Phys. Rev. Lett.* **90**, 107201 (2003).
- [93] Nandy, S., Sharma, G., Taraphder, A. & Tewari, S. Chiral anomaly as the origin of the planar Hall effect in Weyl semimetals. *Phys. Rev. Lett.* **119**, 176804 (2017).
- [94] Ostrovsky, P., Gornyi, I. & Mirlin, A. Interaction-induced criticality in Z_2 topological insulators. *Phys. Rev. Lett.* **105**, 036803 (2010).
- [95] Zhang, Y. & Vishwanath, A. Anomalous aharonov-bohm conductance oscillations from topological insulator surface states. *Phys. Rev. Lett.* **105**, 206601 (2010).
- [96] Bardarson, J. H., Brouwer, P. & Moore, J. Aharonov-bohm oscillations in disordered topological insulator nanowires. *Phys. Rev. Lett.* **105**, 156803 (2010).
- [97] De Juan, F., Ilan, R. & Bardarson, J. H. Robust transport signatures of topological superconductivity in topological insulator nanowires. *Phys. Rev. Lett.* **113**, 107003 (2014).
- [98] Cook, A. & Franz, M. Majorana fermions in a topological-insulator nanowire proximity-coupled to an s-wave superconductor. *Phys. Rev. B* **84**, 201105–201109 (2011).
- [99] Cook, A., Vazifeh, M. & Franz, M. Stability of Majorana fermions in proximity-coupled topological insulator nanowires. *Phys. Rev. B* **86**, 155431 (2012).
- [100] Lutchyn, R. *et al.* Majorana zero modes in superconductor–semiconductor heterostructures. *Nat. Rev. Mat.* **3**, 52–68 (2018).
- [101] Peng, H. *et al.* Aharonov–Bohm interference in topological insulator nanoribbons. *Nat. Mater.* **9**, 225 (2009).
- [102] Xiu, F. *et al.* Manipulating surface states in topological insulator nanoribbons. *Nat. Nanotech.* **6**, 216–221 (2011).
- [103] Hong, S. S., Zhang, Y., Cha, J. J., Qi, X.-L. & Cui, Y. One-dimensional helical transport in topological insulator nanowire interferometers. *Nano Lett.* **14**, 2815–2821 (2014).
- [104] Jauregui, L. A., Pettes, M. T., Rokhinson, L. P., Shi, L. & Chen, Y. P. Magnetic field-induced helical mode and topological transitions in a topological insulator nanoribbon. *Nat. Nano.* **11**, 345–351 (2016).
- [105] Bhattacharyya, B., Sharma, A., Awana, V. P. S., Senguttuvan, T. D. & Husale, S. FIB synthesis of Bi₂Se₃ 1D nanowires demonstrating the co-existence of Shubnikov-de Haas oscillations and linear magnetoresistance. *Journal of Physics Condensed Matter* **29**, 07LT01 (2017).
- [106] Ziegler, J. *et al.* Probing spin helical surface states in topological HgTe nanowires. *Phys. Rev. B* **97**, 035157 (2018).
- [107] Bardarson, J. H. & Moore, J. E. Quantum interference and aharonov–bohm oscillations in topological insulators. *Rep. Prog. Phys.* **76**, 056501 (2013).

References

- [108] Goerbig, M. Electronic properties of graphene in a strong magnetic field. *Rev. Mod. Phys.* **83**, 1193 (2011).
- [109] Abrikosov, A. Quantum linear magnetoresistance. *EPL* **49**, 789 (2000).
- [110] Novak, M., Sasaki, S., Segawa, K. & Ando, Y. Large linear magnetoresistance in the Dirac semimetal TlBiSSe. *Phys. Rev. B* **91**, 041203 (2015).
- [111] Shon, N. H. & Ando, T. Quantum transport in two-dimensional graphite system. *J. Phys. Soc. Jpn.* **67**, 2421–2429 (1998).
- [112] Bena, C. Local density of states in the presence of impurity scattering in graphene at high magnetic field. *Phys. Rev. B* **81**, 045409 (2010).
- [113] Champel, T., Florens, S. & Raikh, M. Nonlocal correlations of the local density of states in disordered quantum Hall systems. *Phys. Rev. B* **83**, 125321 (2011).
- [114] Bindel, J., Ulrich, J., Liebmann, M. & Morgenstern, M. Probing the nodal structure of Landau level wave functions in real space. *Phys. Rev. Lett.* **118**, 016803 (2017).
- [115] Peres, N., Guinea, F. & Neto, A. C. Electronic properties of disordered two-dimensional carbon. *Phys. Rev. B* **73**, 125411 (2006).
- [116] Klier, J., Gornyi, I. & Mirlin, A. Transversal magnetoresistance in Weyl semimetals. *Phys. Rev. B* **92**, 205113 (2015).
- [117] Rensink, M. E. Modification of Friedel oscillations by a magnetic field. *Phys. Rev.* **174**, 744 (1968).
- [118] Horing, N. J. Friedel density oscillations about a Coulombic impurity in a high magnetic field. *Phys. Rev.* **186**, 434 (1969).
- [119] Zhang, C.-L. *et al.* Signatures of the Adler–Bell–Jackiw chiral anomaly in a Weyl fermion semimetal. *Nat. Commun.* **7**, 10735 (2016).
- [120] Jia, S., Xu, S.-Y. & Hasan, M. Z. Weyl semimetals, Fermi arcs and chiral anomalies. *Nat. Mat.* **15**, 1140 (2016).
- [121] Klier, J., Gornyi, I. & Mirlin, A. Transversal magnetoresistance and Shubnikov–de Haas oscillations in Weyl semimetals. *Phys. Rev. B* **96**, 214209 (2017).
- [122] Abrikosov, A. Quantum magnetoresistance. *Phys. Rev. B* **58**, 2788 (1998).
- [123] Shoenberg, D. *Magnetic oscillations in metals* (Cambridge university press, 1984).
- [124] Suzuki, T. *et al.* Singular angular magnetoresistance in a magnetic nodal semimetal. *Science* eaat0348 (2019).
- [125] Zala, G., Narozhny, B. & Aleiner, I. Interaction corrections at intermediate temperatures: Longitudinal conductivity and kinetic equation. *Phys. Rev. B* **64**, 214204 (2001).
- [126] Altshuler, B. L., Aronov, A. G. & Lee, P. Interaction effects in disordered Fermi systems in two dimensions. *Phys. Rev. Lett.* **44**, 1288 (1980).
- [127] Matveev, K., Yue, D. & Glazman, L. Tunneling in one-dimensional non-Luttinger electron liquid. *Phys. Rev. Lett.* **71**, 3351 (1993).

References

- [128] Egger, R. & Grabert, H. Friedel oscillations for interacting fermions in one dimension. *Phys. Rev. Lett.* **75**, 3505 (1995).



Quasi-particle interference from massive Dirac fermions

This appendix calculates the FT-LDOS due to massive Dirac fermions, as given in chapter 3. The procedure for the calculation is largely identical to that of the massless case.

In the presence of a finite mass m the free Green's function has the form

$$G_0(\mathbf{k}, \omega) = \frac{\omega + v_F(k_x \sigma_y - k_y \sigma_x) + m \sigma_z}{(\omega)^2 - v_F^2(k_x^2 + k_y^2) - m^2}. \quad (\text{A.1})$$

As in chapter we assume a diagonal impurity potential such that $\mathbf{T}(\omega) = T(\omega)\mathbb{1}$ and the change in the LDOS can be written

$$\begin{aligned} \delta\rho(\mathbf{q}, \omega) &= -\frac{1}{\pi} \text{Im} \left\{ \text{Tr} \left\{ \int \frac{d\mathbf{k}}{(2\pi)^2} \mathbf{G}_0(\mathbf{k}, \omega) \mathbf{T}(\omega) \mathbf{G}_0(\mathbf{k} - \mathbf{q}, \omega) \right\} \right\} \quad (\text{A.2}) \\ &= -\frac{2U}{\pi} \text{Im} T(\omega) \Lambda(\mathbf{q}, \omega) = -\frac{2U}{\pi} \text{Im} T(\omega) \left\{ \int \frac{d\mathbf{k}}{(2\pi)^2} \frac{\omega^2 + v_f^2 k^2 + m^2 - v_F^2 \mathbf{k} \cdot \mathbf{q}}{(\omega^2 - v_F^2 k^2 - m^2)(\omega^2 - v_F^2 (\mathbf{k} - \mathbf{q})^2 - m^2)} \right\}. \end{aligned}$$

The calculation proceeds identically to the massless case: switching to matsubara frequencies, using the Feynman trick and then integral taking the k integral, which now looks like

$$\int_0^1 dx \int \frac{d^2 \mathbf{k}}{(2\pi)^2} \frac{-\omega^2 + m^2 + v_f^2(\mathbf{k}^2 - x(1-x)\mathbf{q}^2)}{(\Delta^2 + v_F^2 \mathbf{k}^2 + m^2)^2} \quad (\text{A.3})$$

$$= \int_0^1 dx \int_0^\infty \frac{dk}{2\pi} \frac{k}{k} \frac{-\omega^2 + m^2 + v_f^2(k^2 - x(1-x)q^2)}{(\Delta^2 + v_F^2 k^2 + m^2)^2}, \quad (\text{A.4})$$

where $\Delta^2 = \omega^2 + x(1-x)v_F^2 \mathbf{q}^2$ giving

$$\Lambda(q) = \frac{1}{4\pi v_F^2} \int_0^1 dx \left(\frac{-\Delta^2 + m^2}{\Delta^2 + m^2} + \frac{-\Lambda^2 + \Delta^2 + m^2}{\Lambda^2 + \Delta^2 + m^2} + \log \left(\frac{\Delta^2 + m^2}{\Lambda^2 + \Delta^2} \right) \right) \quad (\text{A.5})$$

$$= \frac{1}{4\pi v_F^2} \int_0^1 dx \left(-1 + \frac{-\Delta^2 + m^2}{\Delta^2 + m^2} + \log \left(\frac{\Delta^2 + m^2}{\Lambda^2} \right) \right), \quad (\text{A.6})$$

Appendix A. Quasi-particle interference from massive Dirac fermions

where we used $\Lambda \gg q$ in the final line. This is same as the massless case for $m \rightarrow 0$. Performing the integral gives

$$\Lambda(\mathbf{q}, i\omega) = \frac{1}{2\pi v_F^2} \left(2 \operatorname{sgn}(\tilde{\omega}) g(\tilde{\omega}) \operatorname{arctan} \frac{1}{g(\tilde{\omega})} + \ln \left(\frac{-(\omega^2 + m^2)}{\Lambda^2} \right) + \frac{2 \left(\frac{2m}{qv_F} \right)^2 \operatorname{arctan} \left(\frac{1}{g(\tilde{\omega})} \right)}{g(\tilde{\omega})} \right). \quad (\text{A.7})$$

where $\tilde{\omega}^2 = \left(\frac{2m}{qv_F} \right)^2 + \left(\frac{2\omega}{qv_F} \right)^2$. As before the LDOS difference is simply, after analytic continuation, $\Delta\rho(\mathbf{q}, \omega) - \frac{1}{\pi} \operatorname{Im} T(\omega) \Lambda(\mathbf{q}, \omega)$. We see that this is of the form $\Delta\rho(\mathbf{q}, \omega) = \rho_0(\mathbf{q}, \omega) + \rho_m(\mathbf{q}, \omega)$, where ρ_0 is the same as the massless result with a renormalised ω and $\rho_m(\mathbf{q}, \omega)$ vanishes at for $m \rightarrow 0$ but is sharply peaked when $v_F q/2 \approx \omega$ since here perfect nesting occurs and there is no protection from backscattering (see chapter 3).

Matsubara sum and vertex corrections

In this appendix we provide two of the missing details for the derivation of the conductivity using the the current-current correlator.

Matsubara sum:

This calculation is broadly similar to the one in [66]. The only difference is that there are two poles in the Dirac Green's function for negative and position energies.

The first correlation function that contributes to the conductivity is the bubble diagram:

$$\pi_{\alpha\beta}^{(0)}(i\Omega) = \frac{1}{\beta} \int \frac{d^2\mathbf{k}}{(2\pi)^2} \sum_{\omega_n} \text{Tr} \langle j^\alpha \mathcal{G}(\mathbf{k}, i\omega_n + i\Omega) j^\beta \mathcal{G}(\mathbf{k}, i\omega_n) \rangle. \quad (\text{B.1})$$

The Matsubara summation can be computed as follows:

$$S = \frac{1}{\beta} \sum_{\omega_n} \text{Tr} \langle j^\alpha \mathcal{G}(\mathbf{k}, i\omega_n + i\Omega) j^\beta \mathcal{G}(\mathbf{k}, i\omega_n) \rangle. \quad (\text{B.2})$$

This sum can be calculated from the contour integral along a circle of radius $R \rightarrow \infty$ in the complex plane [66]:

$$\int \frac{dZ}{2\pi i} n_F(Z) \text{Tr} (j^\alpha \mathcal{G}(\mathbf{k}, Z + i\Omega) j^\beta \mathcal{G}(\mathbf{k}, Z)), \quad (\text{B.3})$$

where $n_F(Z)$ is the Fermi-function and the integral equals the sum plus contributions from above and below the branch cuts at the energy eigenvalues $Z = \epsilon^+$, $Z = \epsilon^-$, $Z = \epsilon^+ - i\Omega$, and $Z = \epsilon^- - i\Omega$ (note ϵ^- will be negative but since we integrate along this branch cut that has no implications for the calculation). This differs from the same calculation in [66] due to the positive and negative branches of the Dirac Green's function. Therefore we can write the bare bubble correlator as:

$$\begin{aligned} S = 2i \int \frac{d\epsilon}{2\pi} & \left(n_F(\epsilon) \text{Tr} (j^\alpha \mathcal{G}(\mathbf{k}, \epsilon + i\Omega) j^\beta \left\{ \mathcal{G}(\mathbf{k}, \epsilon + i\delta) - \mathcal{G}(\mathbf{k}, \epsilon - i\delta) \right\} \right. \right. \\ & \left. \left. + n_F(\epsilon - i\Omega) j^\alpha \left\{ \mathcal{G}(\mathbf{k}, \epsilon + i\delta) - \mathcal{G}(\mathbf{k}, \epsilon - i\delta) \right\} j^\beta \mathcal{G}(\mathbf{k}, \epsilon - i\Omega) \right) \right). \end{aligned} \quad (\text{B.4})$$

This means the correlation function can be written as:

$$\begin{aligned} \pi_{\alpha\beta}^{(0)}(i\Omega) = 2i \int \frac{d\epsilon}{2\pi} \int \frac{d^2\mathbf{k}}{(2\pi)^2} & \left(n_F(\epsilon) \text{Tr} (j^\alpha \mathcal{G}(\mathbf{k}, \epsilon + i\Omega) j^\beta \left\{ \mathcal{G}(\mathbf{k}, \epsilon) - \mathcal{G}^\dagger(\mathbf{k}, \epsilon) \right\} \right. \\ & \left. + n_F(\epsilon - i\Omega) j^\alpha \left\{ \mathcal{G}(\mathbf{k}, \epsilon) - \mathcal{G}^\dagger(\mathbf{k}, \epsilon) \right\} j^\beta \mathcal{G}(\mathbf{k}, \epsilon - i\Omega) \right) \right). \end{aligned} \quad (\text{B.5})$$

Appendix B. Matsubara sum and vertex corrections

Performing analytic continuation we set $i\omega \rightarrow \Omega + i\delta$ and can calculate the longitudinal DC conductivity from:

$$\begin{aligned}
\sigma^{\alpha\alpha} &= -\lim_{\Omega \rightarrow 0} \text{Im} \left(\frac{\pi^{\alpha\alpha}(\Omega)}{\Omega} \right) \\
&= -\lim_{\Omega \rightarrow 0} \text{Re} \left\{ 2e^2 v_F^2 \int \frac{d\epsilon}{2\pi} \int \frac{d^2\mathbf{k}}{(2\pi)^2} \text{Tr} \left(\frac{n_F(\epsilon) - n_F(\epsilon + \Omega)}{\Omega} \sigma^\alpha \mathcal{G}(\mathbf{k}, \epsilon + \Omega) \sigma^\alpha \mathcal{G}^\dagger(\mathbf{k}, \epsilon) \right. \right. \\
&\quad \left. \left. + \frac{n_F(\epsilon + \Omega)}{\Omega} \sigma^\alpha \mathcal{G}(\mathbf{k}, \epsilon + \Omega) \sigma^\alpha \mathcal{G}(\mathbf{k}, \epsilon) - \frac{n_F(\epsilon)}{\Omega} \sigma^\alpha \mathcal{G}^\dagger(\mathbf{k}, \epsilon + \Omega) \sigma^\alpha \mathcal{G}^\dagger(\mathbf{k}, \epsilon) \right) \right\}.
\end{aligned} \tag{B.6}$$

Since we take the trace over the matrices, in the limit, the real parts of the final two terms cancel and so they are purely imaginary and do not contribute meaningfully to the conductivity far from the Dirac point. Hence we arrive at the final equation for the DC conductivity:

$$\sigma^{\alpha\alpha} = e^2 v_F^2 \text{Re} \int \frac{d\epsilon}{2\pi} \int \frac{d^2\mathbf{k}}{(2\pi)^2} n'_F(\epsilon) \text{Tr} \{ \sigma^\alpha \mathcal{G}(\mathbf{k}, \epsilon) \sigma^\alpha \mathcal{G}^\dagger(\mathbf{k}, \epsilon) \}. \tag{B.7}$$

Where $n'_F(\epsilon)$ is the derivative of the Fermi-function w.r.t. ϵ . At zero temperature this is just a delta function such that: $n'_F(\epsilon, T = 0) = -\delta(\epsilon - \mu)$ where μ is the chemical potential.

Vertex corrections mapping:

Here we prove the mapping of the components that form the vertex correction to the conductivity, as set out in the main text chapter 4. The first ladder diagram $\pi_{\alpha\beta}^{(1)}(i\Omega)$ written out explicitly in terms of its indices is:

$$\begin{aligned} & \frac{e^2 v_f^2}{\beta} \sum_{\mathbf{k}, \mathbf{k}', \{a_i\}} \sigma_{a1,a2}^\alpha \mathcal{G}_{a2,a3}(\mathbf{k}, i\omega_n + i\Omega) T_{a3,a4}(i\omega_n + i\Omega) \mathcal{G}_{a4,a5}(\mathbf{k}', i\omega_n + i\Omega) \\ & \quad \times \sigma_{a5,a6}^\alpha \mathcal{G}_{a6,a7}(\mathbf{k}', i\omega_n) T_{a7,a8}(i\omega_n) \mathcal{G}_{a8,a1}(\mathbf{k}, i\omega_n) \\ & = \frac{e^2 v_f^2}{\beta} \sigma_{a1,a2}^\alpha M_{a8,a1}^{a2,a3} T_{a7,a8}^{a3,a4} M_{a6,a7}^{a4,a5} \sigma_{a5,a6}^\alpha, \end{aligned} \quad (\text{B.8})$$

where we have defined:

$$M_{a_i,a_j}^{a_p,a_q}(i\omega_n, i\Omega) = \sum_{\mathbf{k}} \mathcal{G}_{a_p,a_q}(\mathbf{k}, i\omega_n + i\Omega) \mathcal{G}_{a_i,a_j}(\mathbf{k}, i\omega_n) \quad (\text{B.9})$$

and

$$T_{a_i,a_j}^{a_p,a_q}(i\omega_n, i\Omega) = T_{a_p,a_q}(i\omega_n + i\Omega) T_{a_i,a_j}(i\omega_n). \quad (\text{B.10})$$

As described in the main text these objects map to 4×4 matrices that are the Kronecker products of the original matrices. Explicitly writing this out in terms of indices we have:

$$\begin{aligned} T_{a_i,a_j}^{a_p,a_q}(i\omega_n, i\Omega) &= T_{a_p,a_q}(i\omega_n + i\Omega) T_{a_i,a_j}(i\omega_n) \\ &\rightarrow T_{a_p,a_q}(i\omega_n + i\Omega) \otimes T_{a_i,a_j}(i\omega_n) = T(i\omega_n + i\Omega)_{2(a_p-1)+a_i, 2(a_q-1)+a_j}. \end{aligned} \quad (\text{B.11})$$

Hence we can write the vertex correction as:

$$\begin{aligned} \pi_{\alpha\beta}^{(0)}(i\Omega) + \pi_{\alpha\beta}^{(1)}(i\Omega) &= \pi_{\alpha\beta}^{(0)}(i\Omega) + \sum_n \pi_{\alpha\beta}^{(n)}(i\Omega) \\ &= \frac{e^2 v_f^2}{\beta} \sum_{i\omega_n} \sigma^\alpha M(i\omega_n, i\Omega) \sum_{n=0}^{\infty} (T(i\omega_n, i\Omega) M(i\omega_n, i\Omega))^n \sigma^\beta, \end{aligned} \quad (\text{B.12})$$

as in the main text. The calculation then proceeds as set out in the main text.



Parameter averaging in the planar Hall effect

This appendix follows from chapter 5 on the planar Hall effect. It describes how to perform the calculation of anisotropic magnetoresistance and density of states, including vertex corrections, averaged over a finite distribution of parameters. We find that our conclusions our calculations using a single parameter are robust.

In reality the topological insulator surface from chapter 5 will contain different types of impurities which are described by a distribution of parameters V and ϵ . To check the robustness of our description of the experiment, we show in the following that such distributions do not affect our conclusions.

Within the self-consistent T-matrix approximation, the average over parameters can directly be implemented by averaging over ϵ and/or V in Eq. (5.12). To perform this averaging we assume that the distribution of parameters is described by a Gaussian distribution $n(\epsilon, V)$ with averages $\bar{\epsilon}$ and \bar{V} , widths Δ_ϵ and Δ_V , and $\int n(\epsilon, V) d\epsilon dV = n^{\text{imp}}$. To calculate vertex corrections, Eq. (5.21) has to be replaced by

$$\mathcal{T}(\omega, \Omega) = \int d\epsilon dV \frac{n(\epsilon, V)}{n^{\text{imp}}} T(\omega + \Omega) \otimes T(\omega). \quad (\text{C.1})$$

Figure C.1 shows the results of both V and ϵ averaging. We see that the distribution has no effect on the position of the peaks in $\delta(\mu) = (\rho_{\parallel} - \rho_{\perp})/\rho_{\perp}$, which are only slightly broadened. The height of the peaks is increased due to the enhanced magnetic scattering arising from impurities with smaller ϵ and V . Most importantly, averaging over impurity distributions does not affect our interpretation of the anisotropic magnetoresistance put forward in the main text.

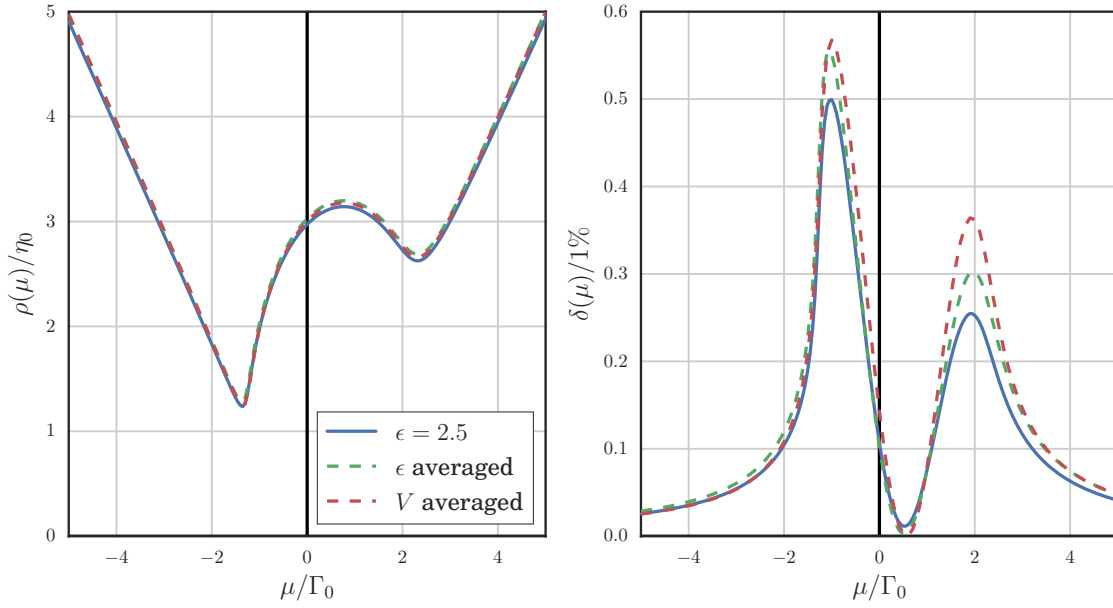


Fig. C.1 Density of states and AMR for averaged parameters: The density of states (left) and resistivity anisotropy $\delta(\mu)$ (right) as a function of chemical potential for a single type of impurity with $\epsilon = 2.5\eta_0$, $V = \sqrt{4\pi}$ (blue line), a Gaussian distribution of ϵ values with $\bar{\epsilon} = 2.5\eta_0$ and width $\Delta_\epsilon = \eta_0$ (green, dashed), and a Gaussian distribution of V values with $\bar{V} = \sqrt{4\pi}$ and width $\Delta_V = \bar{V}/8$ (red, dashed). We see that averaging has a negligible effect on the density of states and on the position of the peaks in $\delta(\mu)$. The peaks are, however, slightly broadened and the size of the peaks increases as impurities with smaller ϵ and smaller V get magnetized more strongly, leading to enhanced spin-flip scattering.

D

3d Dirac Landau level wave-functions

In chapter 7 we have found the Landau level structure of the Dirac Hamiltonian in a magnetic field. We can use this to explicitly determine the wave-functions in real space, as used to calculate the Friedel oscillations in the vicinity of an impurity. The procedure is not dissimilar from that of calculating the eigenfunctions of the Harmonic oscillator. We begin by make use of the lowering operator to write the differential equation

$$\langle x|a|0\rangle = -iv_F \frac{\partial \phi_0}{\partial x} \pm v_F K_{0x} \phi_0 + v_F \frac{\partial \phi_0}{\partial y} - v_F B y \phi_0 = 0. \quad (\text{D.1})$$

Since we know the eigenfunctions of the harmonic oscillator are Gaussians we make use of the ansatz wave function $\phi_0(x, y) = \exp(i(k_x + K_{0x})x) \varphi_0(y)$ which solves this differential equation when $\varphi_0(y) = \exp(-(k_x l_B - y/l_B)^2/2)$. The full ground state wave-function in position space at $\mathbf{r} = (x, y, z)$ is therefore

$$\psi_0(\mathbf{r}) = A_0 \tau_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-(k_x l_B - y/l_B)^2/2} e^{i(k_z \pm K_{0z})z} e^{i(k_x \pm K_{0x})x}, \quad (\text{D.2})$$

with normalisation $A_0 = 1/(\sqrt{l_B} \sqrt{\pi} L_x L_z)$. We chose a gauge with translation symmetry in the x and z direction so unsurprisingly the wave-function in these directions are plane waves. In the y -direction, where translation symmetry is broken by the field, we find that the wave-function in the lowest Landau level decays as a Gaussian centered around $y = k_x$.

The higher Landau levels are easily obtained from the lowest Landau level using the identity $\phi_n = \langle x|n\rangle = \langle x|(a^\dagger)^n|0\rangle/\sqrt{n!}$ which defines the differential equation

$$\begin{aligned} \phi_n &= \frac{1}{\sqrt{n!}} \left(-iv_F \frac{\partial}{\partial x} - v_F K_{0x} - v_F \frac{\partial}{\partial y} - v_F B y \right)^n e^{-(k_x l_B - y/l_B)^2/2} e^{i(k_x + K_{0x})x} \\ &= \frac{1}{\sqrt{n!}} \left(v_F k_x - v_F \frac{\partial}{\partial y} - v_F B y \right)^n e^{-(k_x l_B - y/l_B)^2/2} e^{i(k_x + K_{0x})x}. \end{aligned} \quad (\text{D.3})$$

We recognise the right hand side of this as the same equation as those for the excited states of the Harmonic oscillator, or equivalently, the generating operator for Hermite polynomials. It follows that the full wave functions of the n th LL in real-space

$$\psi_n(k_x; y) = \begin{pmatrix} A_n H_n \left(\frac{y}{l_B} - k_x l_B \right) \\ A_{n-1} H_{n-1} \left(\frac{y}{l_B} - k_x l_B \right) \end{pmatrix} \exp \left\{ -\frac{1}{2} \left(\frac{y}{l_B} - k_x l_B \right)^2 \right\} e^{i(k_x + K_{0x})x} e^{i(k_z + K_{0z})z}, \quad (\text{D.4})$$

where $H_n(\xi)$ is the n th Hermite polynomial and the normalisation is given by $A_n = 1/\sqrt{2^n n! l_B \sqrt{\pi}}$.

Acknowledgements

A PhD thesis may appear from the outside as an individual achievement, in reality it is also the achievement of everyone who supported that individual along the way.

I owe a deep gratitude to my supervisor Prof. Dr. Achim Rosch who put his trust in me as a young and rather “green” master’s student. His love of physics and eagerness to share his knowledge about the subject is not only inspiring but has always served to provide impetus and energy, even when projects have hit snags along the way.

None of the projects in this thesis would be even half as interesting if not supported by excellent experimental collaborators. In particular Prof. Dr. Yoichi Ando and many members (and former members) of his group and the 2nd institute in Cologne, including (but not limited to): Dr. Alexey Taskin, Dr. Mario Novak, Felix Munning, Dr. Oliver Breuing, Dr. Wouter Jolie, and Timo Knispiel.

A special thanks also goes to Dr. Chris Hooley and Dr. Bernd Braunecker at the University of St Andrews for helping me develop as a physicist and academic. Jamie Jones and Dr. Philip Hölzenspies have always kept me intellectually honest and provided great company.

My office mates – in order of appearance – Dr. Jan Gelhausen, Yoran Tournois, Dr. Emilio Torres, Florian Lange, and Ankita Negi have been fantastic friends and have led to lots of lively discussions, sometimes even about physics. The “neighbours,” Dr. Vatsal Dwivedi and Dr. Ciaran Hickey have been equally good discussion partners, mostly over coffee and rugby (respectively) and about physics (collectively). In addition to those already mentioned I’d like to thank Ori Alberton, Dr. Zala Lenarcic, Lorenz Mayer, Dr. Kevin O’Brien, and Federico Tonielli for the great working environment and friendships they have provided in and outside of the physics institute.

I am extremely grateful to Emilio, Lorenz, and Vatsal for reading sections of this thesis and many helpful suggestions and comments.

Johanna Lübke has not only cared for and supported me, but has also helped me improve my German and put up with living with me. (On the latter front so have, at various times and for wildly varying time scales: Yoran, Denny, and Ori).

Last but by far not least my thanks goes to my parents Tony and Elspeth for all their support along the way and for providing a wonderful environment when growing up, where inquisition was encouraged and ideas challenged if unsupported by evidence. In that respect I owe my older brother Edward for always setting me a high standard to follow.

Thank you. Merci bein des fais. Vielen Dank.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbstständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegt hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Achim Rosch betreut worden.

Henry F. Legg
Köln, 2019

Teilpublikationen:

“Planar Hall effect from the surface of topological insulators.” Taskin AA*, **Legg HF***, Yang F, Sasaki S, Kanai Y, Matsumoto K, Rosch A, Ando Y., 2017. *Nature communications*, 8(1), p.1340. (Ref. [70]).

Submitted:

“Quantum-confined surface channels in bulk-insulating topological-insulator nanowires.” Munning F*, Breunig O*, **Legg HF***, Roitsch S, Fan D, Rößler M, Rosch A, Ando Y

**equal contribution*

